*122960*

# STIC-Biotech/ChemLib

**From:**      Whiteman, Brian
**Sent:**      Wednesday, May 26, 2004 7:34 AM
**To:**        STIC-Biotech/ChemLib
**Subject:**   seq search

09/475,704 Barnett et al., 12/30/99

Please search SEQ ID NO: 1, 2, 3, and 4 against us patent and us patent publication databases.

Thank you,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

Searcher:_____
Phone:_____
Location:_____
Date Picked Up:_____
Date Completed:_____
Searcher Prep/Review:_____
Clerical:_____
Online time:_____

TYPE OF SEARCH:
NA Sequences:_____
AA Sequences:_____
Structures:_____
Bibliographic:_____
Litigation:_____
Full text:_____
Patent Family:_____
Other:_____

VENDOR/COST (where applic.)
STN:_____
DIALOG:_____
Questel/Orbit:_____
DRLink:_____
Lexis/Nexis:_____
Sequence Sys.:_____
WWW/Internet:_____
Other (specify):_____

QM nucleic - nucleic search, using sw model

Run on: May 28, 2004, 09:24:06 ; Search time 160.017 Seconds
(without alignments)
5129.286 Million cell updates/sec

Title: US-09-475-704A-3
Perfect score: 1479
Sequence: 1 atgggcgcccgccagcat.........acgaccccctgagcagtaa 1479

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1197.8 | 81.0 | 1515 | 4 | US-09-475-515-4 | Sequence 4, Appli |
| 2 | 1197.8 | 81.0 | 4472 | 4 | US-09-475-515-75 | Sequence 75, Appl |
| 3 | 1197.8 | 81.0 | 4608 | 4 | US-09-475-515-76 | Sequence 76, Appl |
| 4 | 1197.8 | 81.0 | 4689 | 4 | US-09-475-515-74 | Sequence 74, Appl |
| 5 | 1197.8 | 81.0 | 4766 | 4 | US-09-475-515-73 | Sequence 73, Appl |
| 6 | 1195.8 | 80.9 | 2031 | 4 | US-09-475-515-7 | Sequence 7, Appli |
| 7 | 1114.6 | 75.4 | 1853 | 4 | US-09-475-515-5 | Sequence 5, Appli |
| 8 | 1114.6 | 75.4 | 1865 | 4 | US-09-475-515-78 | Sequence 78, Appl |
| 9 | 1114.6 | 75.4 | 1865 | 4 | US-09-475-515-79 | Sequence 79, Appl |
| 10 | 1114.6 | 75.4 | 4319 | 4 | US-09-475-515-6 | Sequence 6, Appli |
| 11 | 1025.2 | 69.3 | 1268 | 4 | US-09-475-515-9 | Sequence 9, Appli |
| 12 | 947.6 | 64.1 | 4307 | 4 | US-09-552-950-2 | Sequence 2, Appli |
| 13 | 939.8 | 63.5 | 9772 | 4 | US-09-552-950-5 | Sequence 5, Appli |
| 14 | 744.4 | 50.3 | 8366 | 4 | US-09-872-733A-6 | Sequence 6, Appli |
| 15 | 738.2 | 49.9 | 1476 | 4 | US-09-184-418C-100 | Sequence 100, App |
| 16 | 738.2 | 49.9 | 8959 | 4 | US-09-184-418C-11 | Sequence 11, Appl |
| 17 | 733.4 | 49.6 | 4338 | 4 | US-09-872-733A-1 | Sequence 1, Appli |
| 18 | 729.8 | 49.3 | 1485 | 4 | US-09-184-418C-73 | Sequence 73, Appl |
| 19 | 729.8 | 49.3 | 9010 | 4 | US-09-184-418C-8 | Sequence 8, Appli |
| 20 | 710.8 | 48.1 | 1486 | 4 | US-09-184-418C-38 | Sequence 38, Appl |
| 21 | 710.8 | 48.1 | 8992 | 4 | US-09-184-418C-82 | Sequence 82, Appl |
| 22 | 701.2 | 47.4 | 1496 | 4 | US-09-184-418C-82 | Sequence 82, Appl |
| 23 | 701.2 | 47.4 | 8972 | 4 | US-09-184-418C-9 | Sequence 9, Appli |
| 24 | 659.8 | 44.6 | 1503 | 4 | US-09-184-418C-29 | Sequence 29, Appl |
| 25 | 659.8 | 44.6 | 8953 | 4 | US-09-184-418C-3 | Sequence 3, Appli |
| 26 | 658.8 | 44.5 | 1479 | 4 | US-09-184-418C-12 | Sequence 12, Appl |
| 27 | 658.8 | 44.5 | 8968 | 4 | US-09-184-418C-1 | Sequence 1, Appli |
| 28 | 656.6 | 44.4 | 4307 | 4 | US-09-184-418C-48 | Sequence 48, Appl |
| 29 | 656.6 | 44.4 | 8966 | 4 | US-09-184-418C-5 | Sequence 5, Appli |
| 30 | 652.6 | 44.1 | 1494 | 4 | US-09-184-418C-47 | Sequence 47, Appl |
| 31 | 650.2 | 44.0 | 4307 | 4 | US-09-55-950-1 | Sequence 1, Appli |
| 32 | 650.2 | 44.0 | 9719 | 4 | US-09-700-304-1 | Sequence 1, Appli |
| 33 | 647 | 43.7 | 3807 | 2 | US-08-417-210A-78 | Sequence 78, Appl |
| 34 | 647 | 43.7 | 3807 | 2 | US-09-136-1599A-78 | Sequence 78, Appl |
| 35 | 645.4 | 43.6 | 7399 | 2 | US-09-418-848A-9 | Sequence 9, Appli |
| 36 | 645.4 | 43.6 | 9709 | 2 | US-08-188-583-5 | Sequence 5, Appli |
| 37 | 645.4 | 43.6 | 9709 | 3 | US-08-388-353-1 | Sequence 1, Appli |
| 38 | 645.4 | 43.6 | 9709 | 4 | US-09-488-551B-1 | Sequence 1, Appli |
| 39 | 645.4 | 43.6 | 9709 | 4 | US-09-309-572-15 | Sequence 15, Appl |
| 40 | 645.4 | 43.6 | 12479 | 4 | US-09-716-096-15 | Sequence 15, Appl |
| 41 | 645.4 | 43.6 | 12494 | 4 | US-09-318-138-13 | Sequence 13, Appl |
| 42 | 645.4 | 43.6 | 12494 | 4 | US-09-935-312-13 | Sequence 13, Appl |
| 43 | 645.4 | 43.6 | 15581 | 3 | US-08-849-760B-33 | Sequence 33, Appl |
| 44 | 645.4 | 43.6 | 15581 | 3 | US-08-646-538-35 | Sequence 35, Appl |
| 45 | 645.4 | 43.6 | 15581 | 3 | US-09-503-222-35 | Sequence 35, Appl |

ALIGNMENTS

RESULT 1
US-09-475-515-4
; Sequence 4, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag
US-09-475-515-4

Query Match 81.0%; Score 1197.8; DB 4; Length 1515;
Best Local Similarity 89.3%;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

```
Qy    1 ATGGGGCGCCCCGGCGCGGCGGCCAGCATCCTGCGCGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC  60
         ||||| |||| ||||||||| |||||||||||||||||| ||||| |||||||||||| |||||
Db    7 ATGGGGCGGCGCCGGCCAGCGTGCTGCAGCCGGGAGAGATCGCG  66

Qy   61 CTCGCGCCCCGGCGGCGCAAGCGAAGTGCTACATGATGAGAGCACACTGGTGTTGGGCCCAGCGCGCGAG  120
         ||| | |||||||||| |||||| ||||| |||||||| |||||||| ||||||||||||||||||||
Db   67 CTCCGCCCCGGCGCAAGAGAAGTACAAGCTGAAGCACATCGTGTGCGCCGCCAGCCGCGAG  126

Qy  121 CTGGAGAAGTTCGCCCTGAACCCCGGCGCCTGCTGGAGAACCAGAGGAGGGCTGCAAGCAGATC  180
         |||||| |||||| |||||||||| |||||| | |||||| |||||| ||||| |||||||||
Db  127 CTGGAGCGCGCTTCGCCGTGAACCCCCGGCGCAGCGAGAGCAGCGAGGGCTTCCGCCAGATC  186

Qy  181 ATCCGCCGCAGCGACACCCCCGCCCTGCGAGACCGGCAGCGAGGAGGTCGAAGAGCGTGTTCAAC  240
         |||| ||||||||||  |||| | |||||| ||||||||||||||  ||| ||||| |||||
Db  187 CTGGGCCAGCCTGCAGCCCAGCCTGCAGACCGGGCAGCGGCCTGTACAAC  246
```

RESULT 2
US-09-475-515-75
; Sequence 75, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
;
US-09-475-515-75
; SEQ ID NO 75
; LENGTH: 4472
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.moduS4.delV1/V2.gag.modSP2
US-09-475-515-75

Query Match       81.0%;  Score 1197.8;  DB 4;  Length 4472;
Best Local Similarity   89.3%;  Pred. No. 2.9e-173;
Matches 1347;  Conservative   0;  Mismatches 132;  Indels 30;  Gaps  4;

Qy       1 ATGGGCGCCCGCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60
           ||||||||||||||||||||||||||||||||||||| ||||||||| ||||||||||||||
Db    2939 ATGGGCGCCCGCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 2998

Qy      61 CTGCGCCCCGGCGGCAAGAAGCAGTACATGCTGAAGCACCTGGTGTGGGCGCCCGAG 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2999 CTGCGCCCCGGCGGCAAGAAGCAGTACAACGGCTGAAGCACATCGTGTGGGCGCCCGAG 3058

Qy     121 CTGGAGAGGTTCGCCGTGAACCCCGGCCCTCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3059 CTGGAGAGGTTCGCCGTGAACCCCGGCCCTCTGGAGACCAGCGAGGGCTGCAAGCAGATC 3118

Qy     181 ATCCGCGAGTTGCACCCCGTGCATGCCGGCCCTGTCGCGGAGCTGAAGAGGCCGTTCAAC 240
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3119 CTGGCCGCCAGCTGCACCCCGTGCACGCCGGAGCTGGCGGAGCTGCGCGACACCAAC 3178

Qy     241 ACCGTGGGCCACCCCGTACTCCGTGCACGAGAAGATCGAGGTCCGCGACAGCCAAGGGAGG 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3179 ACCGTGGGCCACCCCGTGCACGCCGGAGCAGCAGCACCAAGCCGGACGCCCAAGGGAGG 3238

Qy     301 ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGACGTGAAGGACACCAAGGAGGCC 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    3239 CTGGAGCAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGCCGCGGCC 3298

```
Qy        361 GCCG------------ACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 408
              ||||            |||   || |||||||||||||||||||||||||||||||||||
Db       3299 GCCGCCGGCACCGGCAACAGCAGCCAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 3358

Qy        409 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3359 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG 3418

Qy        469 ATCGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGC 528
               | |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       3419 GTGGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGC 3478

Qy        529 GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3479 GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATG 3538

Qy        589 CAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTG 648
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db       3539 CAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTG 3598

Qy        649 CACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3599 CACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC 3658

Qy        709 ACCACCAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTG 768
              |||||||||||||||||||||||||||| ||||||||||| |||||||||||||||||||
Db       3659 ACCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTG 3718

Qy        769 GGCGACATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC 828
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3719 GGCGAGATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC 3778

Qy        829 CCCGTGAGCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGAC 888
              |||   |||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       3779 CCCACCAGCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGAC 3838

Qy        889 CGCTTCTTCAAGACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACC 948
              ||||||| ||||||||||||||| ||||||  || |||||| ||||||||||||||||||
Db       3839 CGCTTCTACAAGACCCTGCGCGCTGAGCAGGCCAGCCAGGACGTGAAGAACTGGATGACC 3898

Qy        949 GACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGC 1008
              || |||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db       3899 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC 3958

Qy       1009 CCCGGCGCCAGCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCAC 1068
              ||||  |||| ||||||||||||||||||||||||||||||||||||||||||| |||||
Db       3959 CCCGCGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAC 4018

Qy       1069 AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGGCCAACA------CCAGCGTGATGATG 1122
              ||||||||||||||||||||||||||||||||||  |  |      | | | | ||||||
Db       4019 AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGGTGACGAACCCGGCGACCATCATGATG 4078

Qy       1123 CAGAAGAGCAACTTCAAGGGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAG 1182
              |||    ||||||||     ||  |||   | ||||||||||||||||||||||||||||
Db       4079 CAGCGCGGCAACTTCCGCAACCAGCGGAAGACCGTCAAGTGCTTCAACTGCGGCAAGGAG 4138

Qy       1183 GGCCACATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAG 1242
              ||||||| |||| | |||||||||||||||||||||||||||||||||   ||||||
Db       4139 GGCCACACCGCCAGGAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGCGCTGCGGCCGC 4198

Qy       1243 GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4199 GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG 4258

Qy       1303 CCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCC 1362
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4259 CCCAGCTACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCC 4318

Qy       1363 CCCGCCGAGAGCTTCCGCTTC------GAGGAGACCACCCCCGGCCAGAAGCAGGAG--- 1413
              ||||  |||||||||||||||      ||| ||||||||||| ||||||||||||||
Db       4319 CCCGAGGAGAGCTTCCGCTTCGGCGAGGAGAAGACCACCCCCAGCCAGAAGCAGGAGCCC 4378
```

```
Qy       1414 ---AGCAAGGACCGCGAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCCCTG 1470
                   |||||| |   |  ||||||||||||||   |||||||||||||||||||||
Db       4379 ATCGACAAGGAGCTGTACCCCCTGACCAGCCTGCGCAGCCTGTTCGGCAACGACCCCAGC 4438

Qy       1471 AGCCAGTAA 1479
              |||||||||
Db       4439 AGCCAGTAA 4447

RESULT 3
US-09-475-515-76
; Sequence 76, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
;  APPLICANT: BARNETT, Susan
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: SRIVASTAVA, Indresh
;  APPLICANT: LIAN, Ying
;  APPLICANT: HARTOG, Karin
;  APPLICANT: LIU, Hong
;  APPLICANT: GREER, Catherine
;  APPLICANT: SELBY, Mark
;  APPLICANT: WALKER, Christopher
;  TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;  TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;  FILE REFERENCE: 1621.002
;  CURRENT APPLICATION NUMBER: US/09/475,515A
;  CURRENT FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 90
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
;   LENGTH: 4608
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: gp160.modSF162.delV2.gag.modSF2
US-09-475-515-76

  Query Match             81.0%;  Score 1197.8;  DB 4;  Length 4608;
  Best Local Similarity   89.3%;  Pred. No. 2.9e-173;
  Matches 1347;  Conservative    0;  Mismatches  132;  Indels   30;  Gaps    4;

Qy          1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60
              |||||||||||||||||| | ||| |||||||| ||||||||   ||||||   ||||||
Db       3075 ATGGGCGCCCGCGCCAGCGTGCTGAGCGGCGGCGAGCTGGACAAGTGGGAGAAGATCCGC 3134

Qy         61 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
              |||||||||||||||||||||   |||| | |||||||| | ||||||||||||||||||
Db       3135 CTGCGCCCCGGCGGCAAGAAGAAGTACAAGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 3194

Qy        121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
              ||||||   |||||| |||||||||||||||||||||||||||||||||||   ||||||
Db       3195 CTGGAGCGCTTCGCCGTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCCGCCAGATC 3254

Qy        181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240
               |  |||||||||| |||  |||||||||||||||||||||||||   ||||||| ||||
Db       3255 CTGGGCCAGCTGCAGCCCAGCCTGCAGACCGGCAGCGAGGAGCTGCGCAGCCTGTACAAC 3314

Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300
              ||||||||||||||||||||||||||| ||   ||||| |||   |||||||||||||||
Db       3315 ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGACGTCAAGGACACCAAGGAGGCC 3374

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
              ||||| ||||||||||||||||||||||||| | || |||||  |||||||||||  |||
Db       3375 CTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGCCGCCGCC 3434

Qy        361 GCCG------------ACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 408
              ||||            |||   || |||||||||||||||||||||||||||||||||||
Db       3435 GCCGCCGGCACCGGCAACAGCAGCCAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 3494

Qy        409 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 4
US-09-475-515-74
; Sequence 74, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 74
; LENGTH: 4689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modSF162.gag.modSF2
US-09-475-515-74

Query Match        81.0%;  Score 1197.8;  DB 4;  Length 4689;
Best Local Similarity   89.3%;  Pred. No. 2.9e-173;
Matches 1347;  Conservative   0;  Mismatches 132;  Indels   30;  Gaps   4;

[The remainder of the page consists of the nucleotide sequence alignment between the Query (Qy) and Database (Db) sequences, shown in paired rows with connecting match lines ("|"), beginning at Qy/Db position 1 and continuing through approximately position 4583, with sequence positions numbered at the margins.]

APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 73
LENGTH: 4766
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: gp160.modUS4.gag.modSF2
US-09-475-515-73

Query Match      81.0%;  Score 1197.8;  DB 4;  Length 4766;
Best Local Similarity   89.3%;  Pred. No. 2.9e-173;
Matches 1347;  Conservative  0;  Mismatches 132;  Indels 30;  Gaps  4;

```
Qy      1 ATGGGCGCCCGCGCCGCAGCATCCTGCGCGGCGGCAAGCTGGACGCGCATCGC 60
          ||||||| |||||||||||||||| ||||| |||||||||||||||||||||||
Db     3233 ATGGGGCGCCGCGCCGCCGCAGCTTCGCGGCGCGGCAAGCTGGACGCGCATCCGC 3292

Qy     61 CTGCGCCCGGCGGCGCAAGAAGTGCTACATGATGAAGCACCTGTGTGGGCGCGAG 120
          |||||| || |||||| || |||||||||||||||||||||| ||||||||||||
Db     3293 CTGCGCCCGGCGCAAGAGAGTACAAGCTGCACTCGTGTGGGCCAGCGCGCGGAG 3352

Qy    121 CTGGAGAAGTTCGCCCTGAAACCCGGCCTCGCTGCTGGAGACCAGCCAGATC 180
          |||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   3353 CTGGAGAAGTTCGCCCTGAAACCCGGCCTCGCTGCTGGAGACCAGCCAGATC 3412

Qy    181 ATCCGGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGCGAGGAGCCCTGTTCAAC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3413 CTGGGCCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGCGAGGAGCCCTTACAAC 3472

Qy    241 ACCGTGGCCACCACCTGTACTGCGTGCACGAAGAGATCGAGGTCCCGGACAGGGCCC 300
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3473 ACCGTGGGCCACCCTGTACTGCGTGCACCAGCCAGCATCGAGGAGGAGGAGGACGGC 3532

Qy    301 CTGGACAAGATCGAGCGAGGCAGCGAGCAGAACAGTGCCAGCAGAAGAATCCAGCGCC 360
          |||| || ||||| |||||||||| || ||||| ||||||||| |||||||||||||
Db   3533 CTGGAGAGATCGAGGCGGAGGAGGAAGTCAAGAGAAGCGGCCCGCGCGCCCCGCC 3592

Qy    361 GCCG----------ACAAGGGCAAGGTGAGCGACGAACTACCCGTGCACAGCTG 408
          ||||          || |||||||||||||||||||||| ||||||||||||||
Db   3593 GCCGCCCCCCAGGACTCGAACACGAACACGATGTTGAACACCGTGGCGGGCGCCCATG 3652

Qy    409 CAGGGGCAGATGTGCACCAGGCCATCAGCCCTCAAACGCCTCGGGTGAAGGTG 468
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3653 CAGGGCCACAGATGTGCACCAGGCCATCAGCCCTCGAACGCCTCGGGTGAAGGTG 3712

Qy    469 ATCGAGGAGAAGGCCTTCAGCCCGAGGTGATCCCATCGTTCACGGCGCCGCC 528
          ||||| |||||||||||||||||||||| |||||||||||||||||||||||
Db   3713 GTGGAGGAGAAGGCCTTCAGCCCGGAGGTGATCCCCATCGTTCACGGCGCCGCC 3772

Qy    529 GCCACCCCCAGGACCTGAACACGATGTTGAACACCGTGGCGGGCGCCACCATG 588
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   3773 GCCACCCCCAGGACCTGAACAACACGATGTTGAACACCGTGGACCGCGCCACCATG 3832

Qy    589 CAGATGCTGAAGGACACACCATCAACGAGGGGCGGCACCCGGTGCACCCGTG 648
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3833 CAGATGCTGAAGGAGACCATCAACGAGGGCGGCACCACCCGGTGCACCCGTG 3892
```

```
Qy    589 CAGATGCTGAAGGAGCACCATCAAACGAGGAGGGCGGCACCCGGTGCACCCCGGTG 648
          ||||| |||||||||| ||||||| |||||| ||||||||||||||||||||||
Db   3696 GCCACCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGGCCACCAGGCCGCCATG 3755
Qy   3756 CAGATGCTGAAGGAGACCATCAACGAGGAGGGCGGCACCCGGTGCACCCGGTG 3815
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3876 ACCACCAGCACCCTGCAGGAGGAGATCGGGTGGATGACCAACAACCCCCCCATCCCCGTG 3935
Qy    769 GGCGACATCTACAAGGGGTTGGATCATCCTGCGGGCCCCTGAAACAAGAGGATCGTGCGATGCTACAGCC 828
          |||||||||||||||||||| |||||||| |||||||||||||||||||||||||||||||||||||||
Db   3936 GGGGAGATCTACAAGGGGTTGGATCATCCTGGGGCCCCTGAAACAAGATCGTGGGATGCTACAGC 3995
Qy    709 ACCAGCAGCACCTGCAGGAGGATGAAGATCGCTGTGATGACCAGCAACCCCCCATCCCCGTG 768
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3816 CACGCGGCGCCCCCATGGCCGCCATCGCCCCGGCCAGATGCGCGGAGGCGGACATCGCCGGC 3875
Qy    649 CACGCGCGGCGCCCCCATGGCCGCCATCGCCCCGGCCAGATGCGCGGAGGCGGACATCGCCGGC 708
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 CCCGTGAGCATCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGAC 888
Qy    829 CCCGTGAGCATCTGGACATCAAGCAGGGCCCCCAAGGAGCCCCTTCCGCGACTACGTGGAC 888
          |||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||||
Db   3996 CCCACGCATCCTGGACATCGCGCCCAAGGAGGCCCTTCGGGACTACGTGGAC 4055
Qy    889 CGCTTCTTCAAGACCCTGCGCGCTGCGCCCGAGCAGAGCAAGCCCCAGGAGGTGGATGACC 948
          ||||||||| |||||||||||||| |||||||||||||| ||||||||||||||||||||
Db   4056 CGCTTCTACAAGACCCTGCGCGCCGAGCAGGCCTCCCAGGAGGTGAAGAACTGGATGACC 4115
Qy    949 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC 1008
          ||||| || |||||| |||||||| ||||||||||||||||||| ||||| ||||||||||
Db   4116 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC 4175
Qy   1009 CCCGGGCCCAGCTGATGAGGCGAGGGGTGGGGCCGCCGCCAGCCAC 1068
          |||||||||||||| ||||| |||||||||||| ||||||||||||
Db   4176 CCCGCGGCCAGCCTGGAGGATGAAACCGCCTGAGGGCGCGGGCGGGCTCTCGGC 4235
Qy   1069 AAGGCCCGGCGTGCTGGCGAGGCGGCCAACA------CCAGCGTGATATG 1122
          ||||| |||||||| ||| ||||||||||||      ||||||||||||
Db   4236 AAGGCCCGGCGTGCTGGCGAGCGGAGCCGACGACATCATGATG 4295
Qy   1123 CAGAAGAGACAACTTCAAGGGCCCACTCGTCACTGGGGAAGGAG 1182
          ||||||||| |||||||||||||||||| ||||||||||||||
Db   4296 CAGCCGGGGCAACTTCGGCAACCAGTCAAGTGCTTCAACACTGGGGAGGAG 4355
Qy   1183 GGCCACATCGGCCCCAACTGCGCGCGCGCCCGCCCCCGCCCCCGCCCCCGCC 1242
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4356 GGCCACAGCCAGAGAACTGCGCGCGGCTGCTTGGCGCGTGGCGCGC 4415
Qy   1243 GAGGGCCCACCGATGAAGGACTGCAACGTTCCGGCAGAAGCAGGAG--- 1302
          ||||||||||||||||||||||||||||||| |||||||||||||
Db   4476 ACCGTGGCCACCACCTGTACTGCGTGCACGAAGAGATCGAGGAGGAGGAGGAC 4535
Qy   4416 GAGGGGCCACCGATGAAGGACTGCAACGTGCACCGGGGACCGCCAAGGAGATCTGG 4475
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4416 CCCAGCCCACCAAGGGGCCGCCGCCGCCAGGCCCCCCAGCCCACCCCCGCC 4475
Qy   1303 CTGGAGAGATCGAGGAGGCGGAGGAACAAGTCAAGAGAAAGCGGCCCGCGCC 1362
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4476 CCCAGCGTACAAGGGCCCGCCAACTTCCGCAAGGCCCCACGCGCCCGCC 4535
Qy   1363 CCCGCGCCGGAGGAGCTTCCGCTTC------GAGGAGGAGCACACCCTGAAAGCCAGGAG 1413
          |||||||||||||||||||||      |||||||||||||| |||||||||
Db   1414 ---AGCAAGGACACGCGCGACACCCTGACCGAGGAGCTGCTGCCCAACGACACCCCTG 1470
Qy   1414 ---AGCAAGGACACGCGCGACACCCTGACCGAGGAGCTGCTGCGCCAACGACCCCCTG 1470
          |||||||||||| |||||||||||||||||||||||||||||||
Db   4536 CCCGAGGAGGAGCTTCGGCTTCGCGAGAAGACACCACCCCCGAGGAGGGG 4595
Qy   4536 CCCGAGGAGGAGCTTCGGCTTCGGGCAGAAGCAGGAGGG 4595
          |||||||||||||||||||| ||||| |||| ||||||
Db    469 ATCGACGAAGGAGCTGTACCCCTGGCGGCAGGCCTGCGCAACCCCAGC 528
Qy   1471 AGCCAGTAA 1479
          |||||||||
Db   4596 ATCGACGAAGGAGCTGTACCCCTGGCGGCAGGCCTGCGCAACCCCAGC 4655
Qy   1471 AGCCAGTAA 1479
Db   4656 AGCCAGTAA 4664
```

```
Qy        649 CACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 CACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC 3952

Qy        709 ACCACCAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTG 768
              |||||||||||||||||||||||||||| ||||||||||| |||||||||||||||||||
Db       3953 ACCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTG 4012

Qy        769 GGCGACATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC 828
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GGCGAGATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC 4072

Qy        829 CCCGTGAGCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGAC 888
              |||   |||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       4073 CCCACCAGCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGAC 4132

Qy        889 CGCTTCTTCAAGACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACC 948
              ||||||| ||||||||||||||| ||||||  || |||||| ||||||||||||||||||
Db       4133 CGCTTCTACAAGACCCTGCGCGCTGAGCAGGCCAGCCAGGACGTGAAGAACTGGATGACC 4192

Qy        949 GACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGC 1008
              || |||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db       4193 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC 4252

Qy       1009 CCCGGCGCCAGCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCAC 1068
              ||||  |||| ||||||||||||||||||||||||||||||||||||||||||| |||||
Db       4253 CCCGCGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAC 4312

Qy       1069 AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGGCCAACA------CCAGCGTGATGATG 1122
              ||||||||||||||||||||||||||||||||||  |  |      | | | | ||||||
Db       4313 AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGGTGACGAACCCGGCGACCATCATGATG 4372

Qy       1123 CAGAAGAGCAACTTCAAGGGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAG 1182
              |||    ||||||||     ||  |||   | ||||||||||||||||||||||||||||
Db       4373 CAGCGCGGCAACTTCCGCAACCAGCGGAAGACCGTCAAGTGCTTCAACTGCGGCAAGGAG 4432

Qy       1183 GGCCACATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAG 1242
              ||||||| |||| | |||||||||||||||||||||||||||||||||   ||||||
Db       4433 GGCCACACCGCCAGGAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGCGCTGCGGCCGC 4492

Qy       1243 GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4493 GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG 4552

Qy       1303 CCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCC 1362
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4553 CCCAGCTACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCC 4612

Qy       1363 CCCGCCGAGAGCTTCCGCTTC------GAGGAGACCACCCCCGGCCAGAAGCAGGAG--- 1413
              ||||  |||||||||||||||      ||| ||||||||||| ||||||||||||||
Db       4613 CCCGAGGAGAGCTTCCGCTTCGGCGAGGAGAAGACCACCCCCAGCCAGAAGCAGGAGCCC 4672

Qy       1414 ---AGCAAGGACCGCGAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCCCTG 1470
                   |||||| |   |  ||||||||||||||   |||||||||||||||||||||
Db       4673 ATCGACAAGGAGCTGTACCCCCTGACCAGCCTGCGCAGCCTGTTCGGCAACGACCCCAGC 4732

Qy       1471 AGCCAGTAA 1479
              |||||||||
Db       4733 AGCCAGTAA 4741

RESULT 6
US-09-475-515-7
; Sequence 7, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
;  APPLICANT: BARNETT, Susan
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: SRIVASTAVA, Indresh
;  APPLICANT: LIAN, Ying
;  APPLICANT: HARTOG, Karin
;  APPLICANT: LIU, Hong
;  APPLICANT: GREER, Catherine
;  APPLICANT: SELBY, Mark
;  APPLICANT: WALKER, Christopher
;  TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;  TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;  FILE REFERENCE: 1621.002
;  CURRENT APPLICATION NUMBER: US/09/475,515A
;  CURRENT FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 90
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 2031
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: HIV-Gag/HCV-core fusion polypeptide
US-09-475-515-7

  Query Match             80.9%;  Score 1195.8;  DB 4;  Length 2031;
  Best Local Similarity   89.3%;  Pred. No. 5.8e-173;
  Matches 1345;  Conservative    0;  Mismatches  132;  Indels   30;  Gaps    4;

Qy          1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60
              |||||||||||||||||| | ||| |||||||| ||||||||   ||||||   ||||||
Db          7 ATGGGCGCCCGCGCCAGCGTGCTGAGCGGCGGCGAGCTGGACAAGTGGGAGAAGATCCGC 66

Qy         61 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
              |||||||||||||||||||||   |||| | |||||||| | ||||||||||||||||||
Db         67 CTGCGCCCCGGCGGCAAGAAGAAGTACAAGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 126

Qy        121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
              ||||||   |||||| |||||||||||||||||||||||||||||||||||   ||||||
Db        127 CTGGAGCGCTTCGCCGTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCCGCCAGATC 186

Qy        181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240
               |  |||||||||| |||  |||||||||||||||||||||||||   ||||||| ||||
Db        187 CTGGGCCAGCTGCAGCCCAGCCTGCAGACCGGCAGCGAGGAGCTGCGCAGCCTGTACAAC 246

Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300
              ||||||||||||||||||||||||||| ||   ||||| |||   |||||||||||||||
Db        247 ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGACGTCAAGGACACCAAGGAGGCC 306

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
              ||||| ||||||||||||||||||||||||| | || |||||  |||||||||||  |||
Db        307 CTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGCCGCCGCC 366

Qy        361 GCCG------------ACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 408
              ||||            |||   || |||||||||||||||||||||||||||||||||||
Db        367 GCCGCCGGCACCGGCAACAGCAGCCAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 426

Qy        409 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG 486

Qy        469 ATCGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGC 528
               | |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        487 GTGGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGC 546

Qy        529 GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATG 606

Qy        589 CAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTG 648
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        607 CAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTG 666

Qy        649 CACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 CACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC 726

Qy        709 ACCACCAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTG 768
              |||||||||||||||||||||||||||| ||||||||||| |||||||||||||||||||
Db        727 ACCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTG 786
```

```
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag-protease
US-09-475-515-5

  Query Match          75.4%;  Score 1114.6;  DB 4;  Length 1853;
  Best Local Similarity   85.8%;  Pred. No. 1.1e-160;
  Matches 1295;  Conservative  0;  Mismatches 184;  Indels  30;  Gaps  4;
```

```
RESULT 7
US-09-475-515-5
; Sequence 5, Application US/0945515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
```

(Sequence alignment block of Qy/Db rows with DNA sequences, positions 1–1853, omitted for legibility.)

RESULT 8
US-09-475-515-78
; Sequence 78, Application US/09475515A
; Patent No. 602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 78
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: GP1
US-09-475-515-78

Query Match     75.4%;  Score 1114.6;  DB 4;  Length 1865;
Best Local Similarity   85.8%;  Pred. No. 1.1e-160;
Matches 1295;  Conservative 0;  Mismatches 184;  Indels 30;  Gaps 4;

(DNA sequence alignment between Query (Qy) and Database (Db), positions 1–1865)

RESULT 9
US-09-475-515-79
; Sequence 79, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 79
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GP2
US-09-475-515-79

Query Match    75.4%;    Score 1114.6;   DB 4;   Length 1865;
Best Local Similarity   85.8%;
Matches 1295;    Conservative    0;   Mismatches 184;   Indels   30;   Gaps   4;

RESULT 10
US-09-475-515-6

; Sequence 6, Application US/09475515A
; Patent No. 662705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 4319
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag-polymerase

Query Match     75.4%;  Score 1114.6;  DB 4;  Length 4319;
Best Local Similarity   85.8%;  Pred. No. 1.1e-160;
Matches 1295;  Conservative   0;  Mismatches 184;  Indels  30;  Gaps   4;

[Paired nucleotide sequence alignment (Qy = query, Db = database subject), positions as follows:]

Qy  121 ... 180          Db  127 ... 186
Qy  181 ... 240          Db  187 ... 246
Qy  241 ... 300          Db  247 ... 306
Qy  301 ... 360          Db  307 ... 366
Qy  361 ... 408          Db  367 ... 426
Qy  409 ... 468          Db  427 ... 486
Qy  469 ... 528          Db  487 ... 546
Qy  529 ... 588          Db  547 ... 606
Qy  589 ... 648          Db  607 ... 666
Qy  649 ... 708          Db  667 ... 726
Qy  709 ... 768          Db  727 ... 786
Qy  769 ... 828          Db  787 ... 846
Qy  829 ... 888          Db  847 ... 906
Qy  889 ... 948          Db  907 ... 966
Qy  949 ... 1008         Db  967 ... 1026
Qy 1009 ... 1068         Db 1027 ... 1086
Qy 1069 ... 1122         Db 1087 ... 1146
Qy 1123 ... 1182         Db 1147 ... 1206
Qy 1183 ... 1242         Db        ... 1242

Qy    1 ATGGGGCGCCCCGCCGCATCCTGCGCGGCGGCCAAGCTGGACGCGTGGGAGCGCATCCGC 60
Db    7 ATGGGCCCCGCCGCCGCCGCCAAGCTGGACGCGCTGGAGCGCGCAGAGATCCGC 66

Qy   61 CTGCGCCCCGGGGGGAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
Db   67 CTGCGCCCCGGCGGCAAGAAGAAGTACAAGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 126

Db 1093 ... 1152         Qy 1123 ... 1182
Db 1153 ... 1212         Qy 1183 ... 1242
Db 1213 ... 1272         Qy 1243 ... 1302
Db 1273 ... 1332         Qy 1303 ... 1362
Db 1303 ... 1366         Qy 1333 ... 1392
Db 1367 ... 1426         Qy 1363 ... 1413
Db 1393 ... 1452         Qy 1393 ... 1452
Db 1453 ... 1470         Qy 1414 ... 1470
Db 1453 ... 1512         Qy 1471 ... 1479
Db 1513 TCACAGTAA 1521   Qy 1513 AGCCAGTAA 1479

```
              ||||||| |||| | ||||||||||||||||||||||||||||||||   ||||||
Db       1207 GGCCACACCGCCAGGAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGCGCTGCGGCCGC 1266

Qy       1243 GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG 1302
              || || ||||| ||||| || ||||| ||| | ||||| || ||  | || |||||||||
Db       1267 GAAGGACACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTTTTTAGGGAAGATCTGG 1326

Qy       1303 CCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCC 1362
              ||   | |||||||  | || || || || || |||||| | || ||||| || |||||
Db       1327 CCTTCCTACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCCCCA 1386

Qy       1363 CCCGCCGAGAGCTTCCGCTTCGAGGAG------ACCACCCCCGGCCAGAAGCAGGAG--- 1413
              || |  ||||||||| | || | ||||      || || |||   ||||||||||||
Db       1387 CCAGAAGAGAGCTTCAGGTTTGGGGAGGAGAAAACAACTCCCTCTCAGAAGCAGGAGCCG 1446

Qy       1414 ---AGCAAGGACCGCGAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCCCTG 1470
                   |||||| |   |  |  | ||   ||| |     || || ||||||||||||  |
Db       1447 ATAGACAAGGAACTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCG 1506

Qy       1471 AGCCAGTAA 1479
                 ||||||
Db       1507 TCACAGTAA 1515

RESULT 11
US-09-475-515-9
; Sequence 9, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
;  APPLICANT: BARNETT, Susan
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: SRIVASTAVA, Indresh
;  APPLICANT: LIAN, Ying
;  APPLICANT: HARTOG, Karin
;  APPLICANT: LIU, Hong
;  APPLICANT: GREER, Catherine
;  APPLICANT: SELBY, Mark
;  APPLICANT: WALKER, Christopher
;  TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;  TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;  FILE REFERENCE: 1621.002
;  CURRENT APPLICATION NUMBER: US/09/475,515A
;  CURRENT FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 90
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 1268
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
;   OTHER INFORMATION: common region
US-09-475-515-9

  Query Match             69.3%;  Score 1025.2;  DB 4;  Length 1268;
  Best Local Similarity   89.6%;  Pred. No. 3.5e-147;
  Matches 1131;  Conservative    0;  Mismatches  113;  Indels   18;  Gaps    2;

Qy          1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60
              |||||||||||||||||| | ||| |||||||| ||||||||   ||||||   ||||||
Db          7 ATGGGCGCCCGCGCCAGCGTGCTGAGCGGCGGCGAGCTGGACAAGTGGGAGAAGATCCGC 66

Qy         61 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
              |||||||||||||||||||||   |||| | |||||||| | ||||||||||||||||||
Db         67 CTGCGCCCCGGCGGCAAGAAGAAGTACAAGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 126

Qy        121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
              ||||||   |||||| |||||||||||||||||||||||||||||||||||   ||||||
Db        127 CTGGAGCGCTTCGCCGTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCCGCCAGATC 186

Qy        181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240
               |  |||||||||| |||  |||||||||||||||||||||||||   ||||||| ||||
Db        187 CTGGGCCAGCTGCAGCCCAGCCTGCAGACCGGCAGCGAGGAGCTGCGCAGCCTGTACAAC 246

Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300
              ||||||||||||||||||||||||||| ||   ||||| |||   |||||||||||||||
Db        247 ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGACGTCAAGGACACCAAGGAGGCC 306

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
              ||||| ||||||||||||||||||||||||| | || |||||  |||||||||||  |||
Db        307 CTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGCCGCCGCC 366

Qy        361 GCCG------------ACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 408
              ||||            |||   || |||||||||||||||||||||||||||||||||||
Db        367 GCCGCCGGCACCGGCAACAGCAGCCAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 426

Qy        409 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG 486

Qy        469 ATCGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGC 528
               | |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        487 GTGGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGC 546

Qy        529 GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATG 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATG 606

Qy        589 CAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTG 648
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        607 CAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTG 666

Qy        649 CACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 CACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC 726

Qy        709 ACCACCAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTG 768
              |||||||||||||||||||||||||||| ||||||||||| |||||||||||||||||||
Db        727 ACCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTG 786

Qy        769 GGCGACATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC 828
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 GGCGAGATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC 846

Qy        829 CCCGTGAGCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGAC 888
              |||   |||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        847 CCCACCAGCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGAC 906

Qy        889 CGCTTCTTCAAGACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACC 948
              ||||||| ||||||||||||||| ||||||  || |||||| ||||||||||||||||||
Db        907 CGCTTCTACAAGACCCTGCGCGCTGAGCAGGCCAGCCAGGACGTGAAGAACTGGATGACC 966

Qy        949 GACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGC 1008
              || |||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        967 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC 1026

Qy       1009 CCCGGCGCCAGCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCAC 1068
              ||||  |||| ||||||||||||||||||||||||||||||||||||||||||| |||||
Db       1027 CCCGCGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAC 1086

Qy       1069 AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGGCCAACA------CCAGCGTGATGATG 1122
              ||||||||||||||||||||||||||||||||||  |  |      | | | | ||||||
Db       1087 AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGGTGACGAACCCGGCGACCATCATGATG 1146

Qy       1123 CAGAAGAGCAACTTCAAGGGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAG 1182
              |||    ||||||||     ||  |||   | ||||||||||||||||||||||||||||
Db       1147 CAGCGCGGCAACTTCCGCAACCAGCGGAAGACCGTCAAGTGCTTCAACTGCGGCAAGGAG 1206

Qy       1183 GGCCACATCGCCCGCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAG 1242
              ||||||| |||| | ||||||||||||||||||||||||||||||||   ||||||
Db       1207 GGCCACACCGCCAGGAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGCGCTGCGGCCGC 1266

Qy       1243 GA 1244
              ||
Db       1267 GA 1268
```

## RESULT 12

US-09-552-950-2
Sequence 2, Application US/09552950
Patent No. 6541248
GENERAL INFORMATION:
APPLICANT: Oxford Biomedica (UK) Limited
TITLE OF INVENTION: Anti-Viral Vectors
FILE REFERENCE: 674524-2004
CURRENT APPLICATION NUMBER: US/09/552,950
CURRENT FILING DATE: 2000-04-20
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4307
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:gagpol-SYNgp - codon
OTHER INFORMATION: optimised gagpol sequence
US-09-552-950-2

Query Match 64.1%; Score 947.6; DB 4; Length 4307;
Best Local Similarity 80.7%; Pred. No. 1.8e-135;
Matches 1135; Conservative 0; Mismatches 259; Indels 12; Gaps 2;

```
Qy      1 ATGGGGCGCCCGGGCGCCAGCATCCTGCGCGGGCGGCCCTGGAGCGGCCATCCGC 60
          |||||||||||||||||||||||||||||||||||||  |||||||||||||||
Db      1 ATGGGGCGCCCGGGCGCCAGCATCCTGCGCGGGCGGCCCTGGAGCGGCCATCCGC 60

Qy     61 CTGCGCGGGCCCGGCGGGGCAAGAAGTGCTACATGAATAGAAGCACTCTGGTTCGGGGAG 120
          |||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db     61 CTGCGCGGGCCCGGCGGGGCAAAAAGAAGTACAAGCATCGTGTGGGGCCACCCGGCGAA 120

Qy    121 CTGGAGAAGTTCGCCCTGAAACCCGGCCTGCTGGAGACCTCCGAAGGCTGCAAGCAGATC 180
          |||  ||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGGAGGCGCTTCGCCGTGAAACCCGGCCTCTTCGGACGACCAAGACCAAGATC 180

Qy    181 ATCCGCGACAGCTCCACCCGCCGCCCGCTGCAGAACCGGCGCCGAGCCGCCTGTACAAC 240
          |||||||||||||||||||  |||||||||||||||||||||||||||||
Db    181 ATCCGCGACAGCTCCACCCGCCGCGGGCGCAGCGGAGCTGAAGAGCCCTGTTCAAC 240

Qy    241 ACCGTGGCCCACCGCTGTACTGGCCGCATCGAAAATCAAGGATACGAAAAGAGGCC 300
          |||||||||||||  |||||||||||||||||
Db    241 ACCGTGGCCCACCGCTGTACTGGCCCGCATCGAAATCAAGGATACGAAAGAGGCC 300

Qy    301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAGCAGAGAATCCAGCAGGCCGGAGGCC 360
          |||  |||||||||||||||||||||||||||||||
Db    301 CTGGATAAAATGTGAAGAGGACAGAATAAGAGCCCAACAGCCGCCCG 360

Qy    361 GCC------GACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGC 414
          |||       |||  ||||||||||||||
Db    361 GACACCGGACACAGCCAACCAGGTCAGCGGTTCAGCGAGAACTACCCCATCGTGCAGAAC 420

Qy    415 CAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGGTGGAA 474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGATGGTGCACCAGGCCATCTCCCCGCACCCTGAACGCCTGGGTGAAGGTGGTGGAA 480

Qy    475 GAGAAGGCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGGAGGCGCCACC 534
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAGAGGCCTTTAGCCCGGAGGTGATACCCATGTTCTCAGCCCTTGTCAGGAGCGCACC 540

Qy    535 CCCCAGGACCTGAACACAATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATG 594
          |||||  ||||||||||  ||||||||||||||||||||||||||||||||||||||
Db    541 CCCAACTTGAACACCATGCTCAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATG 600

Qy    595 CTTGAAGGACACCATCAACGAGAGAAGCACCAGGGCCGGCCCGTGCACCCGTGCACCCC 654
          ||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGAAGGAGACCATCAATGAGGAGGCTGCGGATGGTGTGCATCCGGTGCACCCA 660

Qy    655 GGCCCCCATCGCCCCCGGCCAGATGCGGGAGCCCCGCGGCAGCGACATCGCCGGGAACACC 714
          ||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db    661 GGGCCCCATCGCGACGGGCTGCGAGCACATCGCCGGGAACGACTCAGACACTGGGAAACGACT 720
```

---

```
Qy    715 AGCACCCCTGCCAGAGGCAGCAGATCGGCTCTTGGATGACCAGAACACCCCCGCCATCCC 774
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    715 AGTACCCCTTCAGAAACCAGAATGGCTGGACGATGACAACAACACCACCTGGGTGGGGCGAC 774

Qy    775 ATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTG 834
          ||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db    781 ATCTACAAACGCTGGATCATCCTGGGCCTGAACAAGATCGTGGCCATGTATAGCCCTAC 840

Qy    835 AGCATCCTGGACATCAAGCAGGAGCCCCAAGGAGCCCTTCCCGGACTACCGGGACCTTC 894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AGCATCCTGGACATCCGCCAAGGCCCCAAGGAGCCCTTTCGGGACTACGTGGACCGGTTC 900

Qy    895 TTCAAGACCCTGCGCGCCCAGCAGCAGCACCCCAGGAGGTGAAGAACTGGATGACCGAC 954
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TACAAAACCCTCCGCGCCGAGCAGGCTAGCCAGGAGGTGAAAAACTGGATGACCGAAACC 960

Qy    955 CTGCTCGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTGCCTCGGCCCCG 1014
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CTGCTGGTCCAGAACGCCAACCCGGACTGCAAGACGATCCTGAAGGCCCTGGGCCCAGC 1020

Qy   1015 GCCAGCCTTGGGAGGAGATGATGACCGCCCTGCCAGGGGCGTCGGCGGCGCCACAGGCC 1074
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCTACCCTGGAGGAAATGATGACCGCCTGTCAGGGAGTGGGCGGACCCCGGCCACAAGGCA 1080

Qy   1075 CGGCTGGTGGCCGGAGACGCCAACCCGGACTGAGAGCGGCGTGAGCGGCCAACA------CCAGGCGGA 1128
          |||||||||||||  ||||||||||||||||||||||||||||||  ||||
Db   1081 CGGCTCCTGGCTGAGGCCATGAGCCAGGTGACCAACTCGCGTACCATCATGATGCAGCGC 1140

Qy   1129 AGCAACTTCAAGCGCCCCGCGGCATCGTCAAGTGTCAACTGCCGCCGCCAC 1188
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GGCAACTTCAGGAACCAACGCAAGATCGTCAAGTGCTTCAACTGTGGCAAAGAGGGCCAC 1200

Qy   1189 ATCGCCCGCAACTGCCGGGCGCCCCGCCCCGCAAGAGAGGCTGTGGGAGTGCGGCAAGGAA 1248
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACAGCCCGCAACTGCAGGGCCCCCTAGGAAAAGGCTGCTGCTGGGAAATGCCGGCAAGGC 1260

Qy   1249 CACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGC 1308
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CACCAGATGAAAGACTGTACGGAGAGACAGGCTAATTTTCTGGGAAAGATCTGGCCTTCC 1320

Qy   1309 CACAAGGGCGGCCCGGCAACTTCCTGCAGAGCCGCCCAGCCCCCGCCCCCGCC 1368
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGAGCCGAGCCGAACCCAACAGCCCCGCC 1380

Qy   1369 GAGAGCTTCCGGTTCGAGGAGACCAC 1394
          ||| ||  |||||||||||||||
Db   1381 GCGAGCTTCAGGTTCGGGGGTCGCGAC 1406
```

---

## RESULT 13

US-09-552-950-5
Sequence 5, Application US/09552950
Patent No. 6541248
GENERAL INFORMATION:
APPLICANT: Oxford Biomedica (UK) Limited
TITLE OF INVENTION: Anti-Viral Vectors
FILE REFERENCE: 674524-2004
CURRENT APPLICATION NUMBER: US/09/552,950
CURRENT FILING DATE: 2000-04-20
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 9772
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:pSYNGP
US-09-552-950-5

Query Match 63.5%; Score 939.8; DB 4; Length 9772;
Best Local Similarity 78.6%; Pred. No. 2.7e-134;
Matches 1182; Conservative 0; Mismatches 297; Indels 24; Gaps 4;

```
Qy      1 ATGGGGCGCGGCCCAGCATCCTGCCGCGGCGGCCAAGCTGGAGCGGCGCATCCGC 60
            ||||||||||||||||||||||||||||||||  |||  |||||||||||||||||
Db   2188 ATGGGGCGCGGCCCAGCGTGCTGCGGCTGCTG||||||||||||||||||||| 60

Qy   1108 CTGCGCCCCCGGCGGCGCAAGGAGAGTGCTACATGATGAACGACCTGGGGGCAGCCGAG 1167
            |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   2188 CGCGGTCCTGGCTGAGGCCATGACGCAGGCGTGACCAAACTCCGCTACCATGATCCAGCGC 2247

Qy     61 CTGCGCCCCGGCGGCGCAAAAATGAAGTACAAAGCTGAACCGGCGAGCCGGAGAGATCCGC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1108 CTGCGAGAAGTTCGGCCTGAACCCCGCCTTCGCTGGTGCGAGGACCAGCAGCGGCCTGAA 1167

Qy    121 CTGGAGAAGTTCGGCCTGAACCCCGCCTTCGCTGG 180
            |||||||||||||||||||||||||||||||||||
Db   1228 CTGGAGGCGCTTCGCCGTGAACCCCGGGCTCCTGG 1287

Qy   1168 CTGCGCCCCGGCGGGCAAAAAATGAAGTACAAAGCTGAACCGGCGAGCCGGAGAGATCCGC 1227
            ||| || | |||||||||||||||||||||||||||||||| |||||||||||||||||
Db   2248 GGCCAACTTCGGACCAACGCAAGATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGGCAC 2307

Qy    181 ATCCCGCCCAGCTGCACCCGTCCGCGAGACCCGGGGCTCTGGGAGGACACCAGCAGGCC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1288 CTCCGCCCAACTGCGCCCCGCAAAATCCCGGGGCTCTGGGGAGGAGCGGTTACAC 1347
```

```
Qy   1477 TAA 1479
            |||
Db   2608 TAA 2610
```

Query Match   50.3%;   Score 744.6;   DB 4;   Length 8366;
Best Local Similarity   70.5%;   Pred. No. 8.5e-105;
Matches 1060;   Conservative   0;   Mismatches 419;   Indels 24;   Gaps 4;

```
Qy        121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
              || ||    |||||  | || || |||||| | || ||     || ||||| |  || ||
Db        890 CTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATA 949

Qy        181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240
               |  | ||||| || ||  |||| ||||| ||    ||||||||||       || | ||||
Db        950 CTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAGGAGCTTCGATCACTATACAAC 1009

Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300
              || || || ||||| || || |||||| ||  ||||||| ||   ||||||||||| || 
Db       1010 ACAGTAGCAACCCTCTATTGTGTGCACCAGCGGATCGAGATCAAGGACACCAAGGAAGCT 1069

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
               | ||||||||| ||||| ||||| ||||||| | || |||||  |||||||||| |  ||
Db       1070 TTAGACAAGATAGAGGAAGAGCAAAACAAGTCCAAGAAGAAGGCCCAGCAGGCAGCAGCT 1129

Qy        361 GCC------GACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGC 414
              | |       |||      |||| ||||| || ||||| || ||||||||| | ||||| 
Db       1130 GACACAGGACACAGCAATCAGGTCAGCCAAAATTACCCTATAGTGCAGAACATCCAGGGG 1189

Qy        415 CAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAG 474
              || ||||| || ||||||||    ||   |||  | || || ||||| ||  ||  | || 
Db       1190 CAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCATGGGTAAAAGTAGTAGAA 1249

Qy        475 GAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACC 534
              |||||||| |||||||| || ||||| ||||||||  | ||  |    || || ||||||
Db       1250 GAGAAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTCAGCATTATCAGAAGGAGCCACC 1309

Qy        535 CCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATG 594
              || ||||||||||||||||||||||||||||||| || || || || |||||||| |||
Db       1310 CCACAGGACCTGAACACGATGTTGAACACCGTGGGGGGACATCAAGCAGCCATGCAAATG 1369

Qy        595 CTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCC 654
               | || || |||||||| |||| || || || |||||  | ||||| || ||||| || 
Db       1370 TTAAAAGAGACCATCAATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCATGCA 1429

Qy        655 GGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACC 714
              || || || || || ||||||||| | || ||  | || || ||||| || || || || 
Db       1430 GGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACTACT 1489

Qy        715 AGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGAC 774
              || ||||| ||||| || || |  ||||||||| |  || || || ||||| || || || 
Db       1490 AGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCTATCCCAGTAGGAGAG 1549

Qy        775 ATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTG 834
              ||||||||| ||||||| ||||||||  |||||||||||||| ||||||| |||||
Db       1550 ATCTACAAGAGGTGGATAATCCTGGGATTGAACAAGATCGTGAGGATGTATAGCCCTACC 1609

Qy        835 AGCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTC 894
              ||||| ||||||||| |  || || |||||| |||||  | ||||| || ||||| |||
Db       1610 AGCATTCTGGACATAAGACAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTC 1669

Qy        895 TTCAAGACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACC 954
              |  || || ||  | || |||||     | |||||||| || || |||||||| || |||
Db       1670 TATAAAACTCTAAGAGCTGAGCAAGCTTCACAGGAGGTAAAAAATTGGATGACAGAAACC 1729

Qy        955 CTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGC 1014
               || |||| || || || ||||| || || ||||||||||||   ||||||||||| |
Db       1730 TTGTTGGTCCAAAATGCGAACCCAGATTGTAAGACCATCCTGAAGGCTCTCGGCCCAGCG 1789

Qy       1015 GCCAGCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCC 1074
              || |  || || || |||||||| || || ||||| || || || ||| |||| ||||| 
Db       1790 GCTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCA 1849

Qy       1075 CGCGTGCTGGCCGAGGCGATGAGCCAGGCCAACA------CCAGCGTGATGATGCAGAAG 1128
               | ||  |||||||||||||||||||||  |  |      | | |   ||||||||||
Db       1850 AGAGTTTTGGCCGAGGCGATGAGCCAGGTGACGAACTCGGCGACCATAATGATGCAGAGA 1909

Qy       1129 AGCAACTTCAAGGGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCAC 1188
               ||||||||  |  ||  |||   ||||||||||||||||| || ||||| || || |||
Db       1910 GGCAACTTCCGGAACCAGCGGAAGATCGTCAAGTGCTTCAATTGTGGCAAAGAAGGGCAC 1969

Qy       1189 ATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGC 1248
```

```
              | |||| | |||||||| |||||||| |||||||||||| |||||| || || ||||| ||
Db       1970 ACCGCCAGGAACTGCCGGGCCCCCCGGAAGAAGGGCTGTTGGAAATGTGGAAAGGAAGGA 2029

Qy       1249 CACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGC 1308
              ||||| ||||| || || || ||| | |||||| || ||  | || ||||||||||||   |
Db       2030 CACCAAATGAAAGATTGTACTGAGAGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCC 2089

Qy       1309 CACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCCCCCGCC 1368
               |||||||  | || || || || || |||||| | || ||||| || ||||| || |
Db       2090 TACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAA 2149

Qy       1369 GAGAGCTTCCGCT------TCGAGGAGACCACCCCCGGCCAGAAGCAGGAG------AGC 1416
              ||||||||| | |      | |||   || || |||   ||||||||||||        |
Db       2150 GAGAGCTTCAGGTCTGGGGTAGAGACAACAACTCCCCCTCAGAAGCAGGAGCCGATAGAC 2209

Qy       1417 AAGGACCGCGAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAG 1476
              ||||| |   |  |  | ||   ||| |     || || ||||||||||||  |  | ||
Db       2210 AAGGAACTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAG 2269

Qy       1477 TAA 1479
              |||
Db       2270 TAA 2272

RESULT 15
US-09-184-418C-100
; Sequence 100, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
;  APPLICANT: Hahn, Beatrice
;  APPLICANT:  Gao, Feng
;  APPLICANT:  Shaw, George
;  TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
;  TITLE OF INVENTION:  IMMUNODEFICIENCY VIRUS TYPE 1
;  FILE REFERENCE: D6287
;  CURRENT APPLICATION NUMBER: US/09/184,418C
;  CURRENT FILING DATE:  1999-11-02
;  NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 100
;   LENGTH: 1476
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus type 1
;   FEATURE:
;   OTHER INFORMATION: isolate=94IN476.104; gene=gag
US-09-184-418C-100

  Query Match             49.9%;  Score 738.2;  DB 4;  Length 1476;
  Best Local Similarity   69.2%;  Pred. No. 8e-104;
  Matches 1023;  Conservative    0;  Mismatches  453;  Indels    3;  Gaps    1;

Qy          1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60
              ||||| ||  | ||    ||  |  | || || ||  | |    |||||    || ||
Db          1 ATGGGTGCGAGAGCGTCAATATTAAGAGGGGGAAAATTAGATAGATGGGAAAAAATTCGG 60

Qy         61 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
               |  | || || || |||||    || |||||| || ||| | || ||||| ||| | |||
Db         61 TTAAGGCCAGGGGGAAAGAAACATTATATGATAAAACACTTAGTATGGGCAAGCAGGGAG 120

Qy        121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
              ||||| |  || || || ||| |||||  |    ||    || || || || || || ||
Db        121 CTGGAAAGATTTGCGCTTAACCCTGGCCTTTTAGAGACGTCAGACGGATGTAAACAAATA 180

Qy        181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240
              ||    ||||| || || || ||  |||| || |  ||||| || | |    | ||||||
Db        181 ATAAAACAGCTACATCCAGCTCTTAAGACAGGAACAGAGGAACTTAGGTCATTATTCAAC 240

Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300
              || || || || || || || || || || |    ||| |||| || |||||||||| |||
Db        241 ACAGTAGCAACTCTCTATTGTGTACATGCAGGGATAGAGGTACGAGACACCAAGGAAGCC 300

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
               | ||||||||| ||||| || || |||||  | ||||| || || |||||||||  | |
Db        301 TTAGACAAGATAGAGGAAGAACAAAACAAAAGTCAGCAAAAAATACAGCAGGCAAAAGAG 360
```

Qy  1441 CTGAAGAGCCTGTTCGGGCAACGACCCCCTGAGCCAGTAA 1479
         ||  ||||  |  |||||| |||||||| ||  ||||
Db  1438 CTCAAATCACTCTTTGGGCAGCGACCCCCTTGTCTCAATAA 1476

Search completed: May 28, 2004, 13:26:33
Job time : 171.017 secs

Qy   361 GCCGACAAAGGGCAAGGTGAGCGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGGCCCAGAGATG 420
Db   361 GCTCGAC---GGAAAGGTCAGTGCAGAAAATTATCCCTATAGTGCAAAATCTCCAAAGGGCAAATG 417

Qy   421 GTGCACCAGGCCATCAGCCCCCGGCATCCTGAACGCCCTGGGTGAAGGTGATCGAGGAGAAG 480
Db   418 GTACACCAGCCCCTATCACCTTAGAACTTTGAATGCGTGGGTAAAAGTAATAGAGGAGAAG 477

Qy   481 GCCTTCAGCCCCCGGAGTGATCCCCATGTTCACCGGCCCTGAGCGGAGGCGCGCCACCCCCAG 540
Db   478 GCTCTTTAGCCCTAATAATCACCATGTTCACAGCATTATCAGAAGCCCACCCCTCT 537

Qy   541 GACCTGAAACACGAGTTGAACACGTGGGCGGGGCCCACCAGCCCCATGCAGAATGCTGAAG 600
Db   538 GAATTTAAAACACCCATGTTAAATACAGTGGGGGGGGACATCAAGCAGCCATGCAAATTGTTAAAA 597

Qy   601 GACACCATCAACGAGGAGGCCCGAGTGTGGCACCGCGTGCACCCGGCCCGC 660
Db   598 GATACCATCAACGAGGAGGCTGCCGGATAGATAGAAGTAGTACATCCAGGGGCCT 657

Qy   661 ATCCGCCCCCGCCACAGTGCGCGAGCGCCCCGGGCCAGGCGACATCGCCCGGGCACCACCAGCACC 720
Db   658 AATCCACCCAGCCCAGATGAGACAACCAAGGGGAAGTGATATAGCCACGGAACTACTAGTACC 717

Qy   721 CTGCAGGAGCAGAATCCCCTGGAATGACCAGCAACCCCCATCCCCGTGGGGGACATCTAC 780
Db   718 CTTCAGGAACAAATAGCCATGGATGGATGACAGGTRACCCACCTATTCCAGTGGAAGACAATCTAT 777

Qy   781 AAGCGGTGGGATCATCCTGGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
Db   778 AAAAGATGGGATAATTCTCGGGGTTAAATAAATAGTRAGAGAATGTATAGCCCTGTTGCCATT 837

Qy   841 CTGGACATCAAGCACGGGGCCCCCAAAGGAGCGCCCTTCCCGGACTACGTGGACCCGCTTCTTCAAG 900
Db   838 TTGGACATAAGACAAGGGGCCAAAGGAACCCCTTTAGGGACTAGTGTACGACCGGTTCTTTAAA 897

Qy   901 ACCCTGCGCGCCCGAGCGAGCACCCAGGGAGGTGAAGAACTNGGAATGGACCGACACCCCCTGGCTG 960
Db   898 ACTTTAAGAGCTGAACAAGCTACACRAAGAGTAACATAAGGGTTGGATGACAGACATCTAT 957

Qy   961 GTGCAGAACGCCAACCTGCTCAAGACCATCCTGCGCGGCTCTCCGGCCCCCGGGGCCAGC 1020
Db   958 GTCCAAAATGCAAAACCCAGATTGTGAAGACCAATTTTTAAGGAGCAATTAGGACCCAGGGGCTTCA 1017

Qy  1021 CTGGAGGAGGATGATGAACCGGCCTGCCGCGGGCGTGGCGGGCGGCCCCAGCGCACAGAGGCCCCGGTG 1080
Db  1018 TTAGAAGAAATGGTGACAGCATGTCAAGGAGTGGGGAGGACCTAGCCACAAAGCAAGAGTG 1077

Qy  1081 CTGGCCGAGGCGATGAGGCCAACTGCAGGCCAACAACTAGGCCTGAGATGAGTRGCAGAAGCAACTTCAAG 1140
Db  1078 GACTGCACCGAGCGCCCAGGCCAACTTCCTGGGCAAGAGTCTGGCCCAGCACAAGGCCGC 1137

Qy  1141 GGCCCCCGGCCGCATCGTCAAGTGCTTCAACTCGCGCGGAGGAGGCCGCCACATCGTCCCCGCAC 1200
Db  1138 GGGCCCTAAAAGAATTGTTAAATGCTTCAACTGTGTGGGCAAGGAGCGGCACATAGCCAGGGAAAT 1197

Qy  1201 TGCCGCGCCCCCCCCGGAAGAAGGCTGTCTGAAAGTGCGGCGCGAAGAGGGGCCGCCCACATGTGAAG 1260
Db  1198 TGCAGGGGCCCCTAGAAAAAGAGGCTCTGTGGAAAATGTGGGCAAGAGAGGAACACCCAAATGAAA 1257

Qy  1261 GACTGCACCGAGACGCGCCCAGGCCAACTTCCTGGGCAAGATCTGGCCCTACCAAGGCGCCGC 1320
Db  1258 GACTGTACTGAGACGGCAGGCCTAATTTTTAGGGAAAATTTGGCCTTCCCAACCAAGGGGAGG 1317

Qy  1321 CCCGGCCAACTTCTCTGCAGACGCGCCACCCGCCCCCCCGCCGGAGGCTTCGC 1380
Db  1318 CCAGGGAATTCCTTCAAAACAGGCCCAGACGCCCAACCAGCCCCAGCAGGACGAGCTTCAGG 1377

Qy  1381 TTCGAGGAGGACCACCCCCCGGCCAGGAGCAGGAGAGGAGGACCGGCGGAGACCCTGACCAGC 1440
Db  1378 TTCAAGGAGAGCAACCCCGCTCCGAAGCAGGGAGTCGAAAGACAGGGAAACCCTTTACTTCC 1437

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:         May 28, 2004, 08:01:50 ; Search time 5783.73 Seconds
                                          (without alignments)
                                          7636.275 Million cell updates/sec

Title:          US-09-475-704A-3
Perfect score:  1479
Sequence:       1 acgggcccgccgccagcat..........acgacccctgagccagtaa 1479

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       27513289 seqs, 1491090276 residues

Total number of hits satisfying chosen parameters:    55026578

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :      EST:*
                1: em_estba:*
                2: em_esthum:*
                3: em_estin:*
                4: em_estmu:*
                5: em_estov:*
                6: em_estpl:*
                7: em_estro:*
                8: em_htc:*
                9: gb_est1:*
                10: g5_est2:*
                11: gb_htc:*
                12: gb_est3:*
                13: gb_est4:*
                14: gb_est5:*
                15: em_estfun:*
                16: em_estom:*
                17: em_est_hum:*
                18: em_gss_hum:*
                19: em_gss_inv:*
                20: em_gss_pln:*
                21: em_gss_vrt:*
                22: em_gss_fun:*
                23: em_gss_mam:*
                24: em_gss_mus:*
                25: em_gss_rod:*
                26: em_gss_pro:*
                27: em_gss_phg:*
                28: em_gss_vrl:*
                29: gb_gss1:*
                    gb_gss2:*

SUMMARIES

| Result No. | Score | Query Match Length | DB | ID | Description |
|---|---|---|---|---|---|
| c  1 | 82.6 | 5.6 | 746 | 14 | CK215205 | CK215205 FGAS02715 |
| c  2 | 81.2 | 5.5 | 749 | 13 | BQ744212 | BQ744212 WHE4113 A |
| c  3 | 80   | 5.4 | 1064 | 14 | CK163121 | CK163121 FGAS01573 |
| c  4 | 79.8 | 5.4 | 1126 | 14 | CK211108 | CK211108 FGAS02294 |

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| | | | | |
|---|---|---|---|---|
| 5 | 78 | 5.3 | 807 | 14 | CB673622 OSJNEe08D |
| 6 | 78 | 5.3 | 821 | 14 | CB662794 OSJNEd07C |
| 7 | 78 | 5.3 | 824 | 14 | CB666192 OSJNEd12P |
| 8 | 78 | 5.3 | 851 | 14 | CB455622 OSJNEb07H |
| 9 | 78 | 5.3 | 807 | 14 | CB673355 OSJNEe07N |
| 10 | 76.6 | 5.2 | 1138 | 14 | CK163513 FGAS01614 |
| 11 | 76.6 | 5.2 | 812 | 14 | CB681973 OSJNEf08M |
| 12 | 76.4 | 5.2 | 832 | 14 | CB684834 OSJNEf14L |
| 13 | 76.4 | 5.2 | 759 | 10 | BF259495 HVSMEf001 |
| 14 | 76.2 | 5.2 | 562 | 10 | BF484304 WHE2321_F |
| 15 | 75.6 | 5.1 | 649 | 14 | CB870888 HC15304w |
| 16 | 75.4 | 5.1 | 834 | 14 | CB678700 OSJNEf01C |
| 17 | 75.4 | 5.1 | 840 | 14 | CB673814 OSJNEe08I |
| 18 | 74.8 | 5.1 | 766 | 14 | CB629976 OSIIB06K |
| 19 | 74.8 | 5.1 | 797 | 14 | CB658304 OSJNEc14E |
| 20 | 74.8 | 5.1 | 810 | 14 | CB618374 OSIIBa02D |
| 21 | 74.8 | 5.1 | 823 | 14 | CB627740 OSIIB020 |
| 22 | 74.8 | 5.1 | 826 | 14 | CB641789 OSJNEb01J |
| 23 | 74.8 | 5.1 | 838 | 14 | CB559169 OSJNCc15O |
| 24 | 74.6 | 5.1 | 865 | 14 | CB628660 OSIIB04F |
| 25 | 74.6 | 5.0 | 708 | 14 | CD878344 AZO4.102I |
| 26 | 74.6 | 5.0 | 785 | 12 | BG369139 HVSMEi002 |
| 27 | 74.2 | 5.0 | 818 | 14 | CD878344 OSJNC02G |
| 28 | 74.2 | 5.0 | 933 | 28 | CC391778 PUHII70TB |
| 29 | 73.6 | 5.0 | 791 | 14 | CB550382 OSJNEb14M |
| 30 | 73.4 | 5.0 | 791 | 14 | CB550382 OSJNEb14M |
| 31 | 72.6 | 4.9 | 2598 | 11 | AY103647 Zea mays |
| 32 | 72.4 | 4.9 | 648 | 10 | BE517305 WHE6616_A |
| 33 | 72.2 | 4.9 | 803 | 14 | CB633391 OSIIEb12D |
| 34 | 72.2 | 4.9 | 1308 | 11 | AY104577 Zea mays |
| 35 | 71.8 | 4.9 | 764 | 13 | BQ804819 WHE3559_C |
| 36 | 71.8 | 4.8 | 718 | 14 | BQ804819 WHE3559_C |
| 37 | 71.4 | 4.8 | 718 | 14 | CD938268 OV.109I23 |
| 38 | 71.2 | 4.8 | 782 | 14 | CD938289 OV.109K04 |
| 39 | 71 | 4.8 | 787 | 14 | CB661708 OSJNEd04P |
| 40 | 71 | 4.8 | 935 | 14 | CB658685 OSJNEc140 |
| 41 | 70.8 | 4.8 | 2299 | 11 | CA280077 SCVPLB2C0 |
| 42 | 70.6 | 4.8 | 738 | 13 | AY106831 Zea mays |
| 43 | 70.6 | 4.8 | 887 | 14 | BU295509 60360B565 |
| 44 | 70.6 | 4.8 | 1146 | 11 | AY104406 Zea mays |
| 45 | 70.4 | 4.8 | 784 | 14 | CB650686 OSJNEd15E |

ALIGNMENTS

RESULT 1
CK215205/c                                      746 bp    mRNA    linear    EST 09-DEC-2003
LOCUS       CK215205                            746 bp    mRNA    linear    EST 09-DEC-2003
DEFINITION  FGAS027158 Triticum aestivum FGAS: Library 6 CAP GATE 1 Triticum
            aestivum cDNA, mRNA sequence.
ACCESSION   CK215205
VERSION     CK215205.1  GI:39621309
KEYWORDS    EST.
SOURCE      Triticum aestivum
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 746)
  AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Budes,P., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
            Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,A.,
            Penniket,C., Roach,J.L. and Sarhan,P.
  TITLE     Functional Genomics of Abiotic Stress In Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033

Email: fgas_ests@cs.usask.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [13,662].

```
FEATURES
source
        Plate: L6B016 row: K column: 15.
        Location/Qualifiers
        1..746
        /organism="Triticum aestivum"
        /mol_type="mRNA"
        /db_xref="taxon:4565"
        /clone_lib="Triticum aestivum FGAS: Library 6 CAP GATE 1"
        /note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Crown
        (50%) and leaf (50%) tissues from wheat cultivar Norstar
        after short exposure times to low temperature in the light
        and in the dark. 12 mRNA populations were combined before
        constructing the library. The first 6 populations: After 7
        days of growth at 20C8 from wheat cultivar Norstar after
        short exposure times to low temperature in the light and
        in the dark. 12 mRNA populations were combined before
        constructing the library. The first 6 populations: After 7
        days of growth at 20, wheat plants were transferred to 4C
        in the light. 1cm crown sections and green leaf tissu were
        separately harvested after 1,3, and 6 hours of low
        temperature exposure. The last 6 populations: After 7 days
        of growth at 20C. wheat plants were transferred to 4C in
        the dark. 1cm crown sections and green leaf tissu were
        separately harvested after 1, 3, and 6 hours of low
        temperature exposure. First strand synthesis in this
        library was done in the presence of methylated dCTP
        thereby protecting from internal cleavage with NotI. In
        addition, this library used a primer for second strand
        synthesis that annealed to an artificial sequence (RNA
        oligo) added before first strand synthesis. Therefore when
        sequences from BST generated from this library will be
        masked for vector and adaptor sequences, an additional
        masking step will have to be included to mask this RNA
        oligo that is common to all clones (sequence
        CGACTGGAGCACGAGGAACACTGACCATGGACTGAAGGAGTAGAAA)."
```

ORIGIN

Query Match
Best Local Similarity   48.4%;
Matches 288; Conservative 0; Pred. No. 0.62; Score 82.6; DB 14; Length 746;
Mismatches 301; Indels 6; Gaps 2;

```
Qy 124 GAGAAGTTCGCCCTGAACCCGGCCTGCTGGAGGGCTGAAGCAGATCATC 183
       |||  |          || | |    ||||  ||| |    ||| |||||
Db 737 GAGATNNCCGAAAAACACCCATGGAACCTGATCGACGTCGCGCGAGCAGGTTC 678

Qy 184 CGCCAGCTGCACCCCGCCCGCTGCAGACCCGGCAGCGGAGAGCTGAAGAGCCTGTTCAACACC 243
       |||  ||  | ||  |   |||| |    |||| ||||||||| |||||||||||| |||||||
Db 677 CTCAAACATGCTCGCTCAAGCTCATCGGGCGCCAAGAGAAGACCATGGAGAGATCGGCGTTCTACACC 618

Qy 244 GTGGGCCACCCTGTACTCGCGTCGCAGAAGAATCGGAGGTCCGCGACACCAAGAGAGGCCCTG 303
       ||||||| ||| |  | || | |||| ||||| ||| || || ||||||| |||||||||| |
Db 617 GGGTACTCCTCGTCGCCACCCGGCTCGCCATCCCGACGACGGCCATCTTGGCCATG 558

Qy 304 GACAAGAGATCGAGGAGGAGCAGAACAGAGTCCACGAGAAGATCCAGGCGAGGCCGCC 363
       ||||| | ||||| || || ||||| | ||| |||||||||||||| || |||||||
Db 557 GACATCAACCGGAGAACTACCAGCTCGGGCTTCCGTCCATCGAGAAGGCCCGGCCGCGTGGG 498

Qy 364 GACAAGGGCAAGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTG 423
       ||||| |||||| |||| | |||  ||  |||| |||| ||||| ||||||| |||| |
Db 497 CACAAGAATCGACHTTCCCGCGAGGG---CCCGGCGCCTGGCCGTGGACGCGCGTGCTGGAG 441

Qy 424 CACCAGGCCCATCAAGCCCGCACCCTCAAACGCTCGGGGTGAAGAGTGATCGGAGGAAGGCC 483
       ||||||||||  |  ||| ||||  | ||| ||  ||| || ||| || |||||| ||||||
Db 440 GACGAGGCCAACCACGGCCACCTTCGACTTCGTCTTCGTGGACGGCGACAAGGACAACTAC 381

Qy 484 TTCAGCCCCGAGGTGATCCCCATGTTCACCGGCGCCCTGAGCGGAGGGAGGGGGCCCCAGGAC 543
       |||  |  ||| ||  ||  ||| | |  |||||||| ||| || || ||| | |||||||||
Db 380 CTCAACTACCACGAGCGCCCTCAGAAGCTGCGTCAAGCTCGGCGGCGGCGCCTCCTCGCTACGAC 321
```

---

RESULT 2
BQ744212

```
LOCUS       BQ744212       749 bp     mRNA     linear     EST 17-JUL-2002
DEFINITION  WHB4113_A01 A01ZS wheat salt-stressed root cDNA library Triticum
            aestivum cDNA clone WHB4113_A01_A01, mRNA sequence.
ACCESSION   BQ744212
VERSION     BQ744212.1  GI:21890999
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
ORGANISM    Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 749)
AUTHORS     Anderson,O.D., Akhunov,E., Chao,S., Crossman,C., Deal,K.,
            Dvorak,J., Lazo,G.R., Pham,J., Rausch,C.J., Wilson,C. and Woo,J.
TITLE       The structure and function of the expressed portion of the wheat
            genomes - Salt-stressed root cDNA library
JOURNAL     Unpublished (2002)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oanders@pw.usda.gov
            Sequences have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: SK primer.
FEATURES             Location/Qualifiers
source          1..749
                /organism="Triticum aestivum"
                /mol_type="mRNA"
                /cultivar="Chinese Spring"
                /db_xref="taxon:4565"
                /clone="WHB4113_A01_A01"
                /tissue_type="Roots"
                /dev_stage="Full tillering"
                /lab_host="E. coli SOLR"
                /clone_lib="wheat salt-stressed root cDNA library"
                /note="Vector: Lambda Uni-ZAP XR, excised phagemid
                pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Hydroponic
                plants grown to full tillering stage were treated with 150
                mM NaCl for either 12 hours or 7 days. Root tissues of the
                plants subjected to both types of treatment were collected
                separately at University of California, Davis (E. Akhunov
                and K. Deal in J. Dvorak's Lab). Total RNA was prepared
                separately from the two samples (12h and 7day treatments),
                and equal amount of RNA was then pooled. PolyA RNA was
                purified from the pooled RNA, a cDNA library was made, and
                the cDNA clones were in vivo excised to give pBluescript
                SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak)
                at the University of California, Davis. Colony plating,
                plasmid DNA preparations and DNA sequencing were performed
                in the OD Anderson lab (all other authors)."
ORIGIN
```

Query Match
Best Local Similarity   47.5%;
Pred. No. 0.93; Score 81.2; DB 13; Length 749;

```
Qy 544 CTGAACACGATGTTGAACACGGTGGGCGGGGGCGGCCACCAGGCGGCCCCATGCAGATGGCTGAAGGAC 603
       |||  || |  ||| |||| |  |     | |||| || ||||| || |   |  | ||   | || | |
Db 320 ---AACACGCCTTCGGAACGGTCCGTGCTCTCCGCGGCGAGGAGCCCG                   264

Qy 604 ACCATCAACGAGGAGGACCCCGGACGTGGGACACCGGCGTGCACCCCGTGCACGCGCGCCCCATC 663
       |  | | | | ||| ||| |  ||  ||  | || |  || || |  |  | || |   | |||
Db 263 ATCCGCTACTACCCGACTTCGTCCTCGACCTCAACAGGCCCTCGCCCGTCGGCGGACCAGCGC 204

Qy 664 GCCCCCGGCCGGCCAGGATGCGGCCGAGCCCCGGGCCAGCCGACGACATCGCCCGGCACCACCAGCA 718
       | |  |  || || ||  |   ||| |  ||| || || ||| |    |||| |  |||  |||  |
Db 203 GTCGAGATCTGCCAGCTCCCCGTCGGACGGGCATCACCCCTCTCCGCGCCGCCGGCCA 149
```

Matches 282; Conservative 0; Mismatches 303; Indels 9; Gaps 1;

```
Qy    280 GTCCCCGACACCAAGAGAGGCCCTTGAGACAGATCGAGGAGGAGGAGAGGAACAAGTGCCAGCAG 339
          ||||||||||||||||||||||||||||||||||||  ||| |||| |||||||||| ||
Db     19 GTCGACGCCAAGCCGCGGCCCCAAGCACCACCGCCGCTTCGGGCCCAAGTGGGGCGTTCAACCAG 78

Qy    340 AAGATCCAGAGCCAGGCTCCGAAGGCCGGAGGCCCCCGACCAAGGCGCAAGGAACTACCCCATCGTG 399
          || ||||  ||||||||||||||| || |||||  |||||| |||||  ||||||| |||
Db     79 CGGCCCCAGCTCCCCCAGGCGGCGACGAGGAGGGCGCCCAAGGGCCCCAAGCCGCCAGGGCGAG 138

Qy    400 CAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATCAGCGCGCCCTTGAACGCCTGG 459
          ||||  ||||||||||||||||  |||     ||||||||||||||||| ||||||||
Db    139 AAGGAGCGGCGCGCCCGCGCCGAGGAGGCCCACTACACGACACCCACCGCCCAGGCC 198

Qy    460 GTGAGAGGTGATCGGAGGAGGAGGAGGGCCCTTCAGCCCGAGGTGATCCCATGTTCACCGCCCTG 519
          || ||||||||| |||||   ||  |||||||||||||||
Db    199 TTCCGCGGGCAACCAGCAGAACACCGGCCTCCGTCGACGATCCAGCCGCACTGGACC 258

Qy    520 AGGGAGGCGGGCCCACACCCCCAGGGACCTGAACCAGTGTTGAACACCGTCGGGGGCCCACCAG 579
          |||||||||||||     |||||||||||||| ||||||||||||||||||||||  ||||||
Db    259 ATCCTCGAGCAGATCCCTTCGCCAACTTCACCAAGCTCTCCTTCGCCGTGCCGACCAG 318

Qy    580 GCCGCCATGACGAGATGCTGAAGGACACCATCAACGAGGAGGGCGCGCCGAAGTGGGACGGCGTG 639
          ||||  |||||||| ||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    319 CCCGGAGGACCTGCTCGTCTGCGGCCCGTACGACGTCGTACGACCCGGCTACGACGCGGTC 378

Qy    640 CACCCCCGTGCACGCGCGCCCCCATGCGCCCCAGATGCGGCCCCGAGGCCCCCGGGGGCCCACGCGAC 699
          |||||  |||||| |||||| ||||||||||||||||||| |||||
Db    379 AACCCCAAGACGGGCCCCGGCTTCAAGAAGTTCAAGATCACC 438

Qy    700 ATCGGCGGCACCACCAGCACCCTGC-----------AGGAGCAGATGCCTGATGGATGACCCAGC 750
          ||||||||||||||||||||||||||           |  || ||| ||||||| ||||
Db    439 ACAACCGACGACCCCATCATCGCGCCAGGAGGAGGACAGAAGGCCCACCGTTCTTGCC 498

Qy    751 AACCCCCCCATCCCCCGTGGGGGACCATCTACAAGGCCGGTGGGATCATCCTGGGCCTGAACAAG 810
          ||||||||| ||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    499 ACCGACAGCCATCTCGCGGCCCTCAACGTGCACGCCCCAAGAAGACCATGAAGATACGGCGTCTACACC 558

Qy    811 ATCGTGCGGAGTGTACAGCCATCCAGTCAAGGCATGTCCTGGACATTCCTGGACATCAAGGAGGCCCTG 864
          ||||||||||||| |||||||||  |||||||||||| |||
Db    559 GTCGTGCAGCGGTCGGCAACAAGCTTCTTCTTCGACAAGACGCGAGGGCTCCCAG 612
```

RESULT 3
LOCUS        CK163121        1064 bp    mRNA    linear    EST 05-DEC-2003
DEFINITION   CK163121 PGAS015739 Triticum aestivum PGAS: Library 4 Gate 8 Tritium
             aestivum cDNA, mRNA sequence.
ACCESSION    CK163121
VERSION      CK163121.1    GI:38993028
KEYWORDS     EST.
SOURCE       Triticum aestivum
  ORGANISM   Triticum aestivum
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Pooideae; Triticeae; Triticum.
REFERENCE    1  (bases 1 to 1064)
  AUTHORS    Allard,P., Crosby,W.L., Danyluk,J., Budes,F., Frick,M., Gaudet,D.,
             Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
             Links,M.G., McCarthy,B.L., Monroy,A., Muzak,I., Nilson,D.,
             Pennikett,C., Roach,J.L. and Sarhan,F.
  TITLE      Functional Genomics of Abiotic Stress in Wheat and Canola Crops
  JOURNAL    Unpublished (2003)
COMMENT      Contact: Wm L Crosby
             Bioinformatics
             University of Saskatchewan, Department of Computer Science
             1C101 Engineering Building, 57 Campus Drive, Saskatoon,
             Saskatchewan, S7N 5A9, Canada
             Tel: 306 966 1769
             Fax: 306 966 2033
             Email: fgas.estsecs.usask.ca
             This sequence is the direct result of the Base calling software
             Phred (default parameters). It is the raw base calls.To aid in the

identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [1,834].

FEATURES             Location/Qualifiers
     source          1..1064
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /clone_lib="Triticum aestivum PGAS: Library 4 Gate 8"
                     /note="Organ: crown and leaf; Vector: pCMV.SPORT6; Aerial
                     parts (crown and leaf) of wheat cultivar Norstar during
                     dehydration stress. 8 mRNA populations were combined
                     before constructing the library. The first four come from
                     removing plants from vermiculite (7 day old plants) and
                     incubating them at 20C on the bench without water for 1,
                     2, 3 and 4 days. The last four come from plants grown in
                     soil in a growth chamber after watering is terminated.
                     Four samplings were taken in a two week period; the first
                     after wilting was observed and the last, two weeks later,
                     consisted of live crown and stem tissue (leaf tissue was
                     yellow and dead). First strand synthesis in this library
                     was done in the presence of methylated dCTP thereby
                     protecting from internal cleavage with NotI."
                     /db_xref="taxon:4565"
                     Plate: L4B007 row: C column: 11.

ORIGIN
     Query Match     5.4%; Score 80; DB 14; Length 1064;
     Best Local Similarity    48.4%;
     Matches 283; Conservative 0; Mismatches 296; Indels 6; Gaps 2;
                     Pred: No. 1,4;

```
Qy    124 GAGAAGTTCGCCCTGAACCCCCGGCCTCTGGAGACCAGCGAGGCTGCAAGCAGATCATC 183
          ||||   |||   | ||||||  ||||||||||||||||
Db    282 GAGATCACCGCCAACCACCCATCGAGACCTGATCACGACGGCTCGGCGGACCAGTTC 341

Qy    184 CGGCCAGCTGCACCCCCGCCCTGCAGACCCGGCGAGGAGCTGTTCAACACC 243
          ||||||  ||||||   |||||||||||||||
Db    342 CTCAACAATGCTGCTCAAGCTCATCCGGCCCAAGAAGACCATGAAGATACGGCGTCTACACC 401

Qy    244 GTGCCCACCCTGTTACTGCCTGCACGAAGAGATCGAGGTCGCGACCACCAAGGAGGCCCTG 303
          ||  |||||||  |||||  |||||||||||||| ||| |||
Db    402 GGCTACTCCCTCCTCGGCCACCGGCCCATCCCAGACGACGGCCACCATCTTGGCCATG 461

Qy    304 GACAAGAATCGAGGAGGAGGAGGAGAACAAGTGCCAGCAGAAGATCCAGCAGCCCGAGGCCGCC 363
          ||||||  ||||||||||| ||||||||||||||||||  ||||
Db    462 GACATCAACCGCGAGAACTACGAGCTCGGGCTCCGCGTGCGTCCGGTCGAGGAGAAGCGCTGGCG 521

Qy    364 GACAAGGGCAAGGTGAGCCAGAACTACCCCATACGTGCAGAACCTGCAGGGCCAGATGGTG 423
          ||||| | ||| |||| |||||  || |  |||||||||||
Db    522 CACAAGAATCGACTTCCCGGAGGG-----CCCGGGCCTGCCGGTGCTGGGTGGCTCGGAG 578

Qy    424 CACCAGGCCATCAGCCCCGCGACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCC 483
          |||||||||||| |||||||||||||||||||||  |||||||||||
Db    579 GACGAGGGCCAACCACGGGCACCTTCGGACTTCGGTGGACGCCAGCAGCAGACAACTAC 638

Qy    484 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGGGAGGCGGCGCCACCCCCAGGAC 543
          |||||| ||||||||||||||||||||||||||
Db    639 CTCAACTACCACGAGCGCTCATGAAGCTCGGTCAAGCTCGGCGGCCGCTCGGCGTACGAC 698

Qy    544 CTGAACACGAGTGTGAACACCGTGGGCGGCGGGCCACCAGGCCGCCATGCAGATGCTGAAGGAC 603
          ||||||||  |||||   |||||||||||||||||  |||
Db    699 ---AAACAGCCTCTGGAACAGCGTCCGTCGTGCTCCCCGGCGACCCCATGCGCAAGTAC 755

Qy    604 ACCATCAACGAGGAGAGGCCGCCCGCCGGTGGGACCCGGTGCACCGCCGCCCATC 663
          ||||||| ||||||| ||| |||| |||||   |||||| ||||||||||
Db    756 ATCCGCTPACTACCGGCGACTTCGTCCTCGACCTCAACACAGCGCCCGGNCGACCAGGCC 815

Qy    664 GCCCCCCGGCCAGATGCGCGAGCCCCCGCGCCCCAGCGACCATCGCCGC 708
          |||||||||||||||   | |||||||||||||||||||||
Db    816 GTCGAGGATCTGCCAGCTCCCCGTCGGCGCGACGGCATCACCCTCTGC 860
```

RESULT 4
CK211108/c

```
LOCUS       CK211108                1126 bp    mRNA    linear   EST 09-DEC-2003
DEFINITION  FGAS022942 Triticum aestivum FGAS: Library 6 CAP GATE 1 Triticum
            aestivum cDNA, mRNA sequence.
ACCESSION   CK211108
VERSION     CK211108.1  GI:39617217
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 1126)
  AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Budes,P., Frick,M., Gaudet,D.,
            Genewein,B., Graf,R., Gulick,P., Hrycan,L.P., Laroche,A.,
            Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE     Functional Genomics of Abiotic Stress In Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas_ests@cs.usask.ca
            This sequence is the direct result of the Base calling software
            phred (default parameters). It is the raw base calls. To aid in the
            identification of the high quality insert the software Lucy
            (default parameters) has been run on this sequence. Lucy identified
            the region [12,750].
            Plate: L6B001 row: A column: 20.
FEATURES             Location/Qualifiers
     source          1..1126
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone_lib="Triticum aestivum FGAS: Library 6 CAP GATE 1"
                     /note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Crown
                     (50%) and leaf (50%) tissues from wheat cultivar Norstar
                     after short exposure times to low temperature in the light
                     and in the dark. 12 mRNA populations were combined before
                     constructing the library. The first 6 populations: After 7
                     days of growth at 20, wheat plants were transferred to 4C
                     in the light. 1cm crown sections and green leaf tissue were
                     separately harvested after 1, 3, and 6 hours of low
                     temperature exposure. The last 6 populations: After 7 days
                     of growth at 20C, wheat plants were transferred to 4C in
                     the dark. 1cm crown sections and green leaf tissue were
                     separately harvested after 1, 3, and 6 hours of low
                     temperature exposure. First strand synthesis in this
                     library was done in the presence of methylated dCTP
                     thereby protecting from internal cleavage with NotI. In
                     addition, this library used a primer for second strand
                     synthesis that annealed to an artificial sequence (RNA
                     oligo) added before first strand synthesis. Therefore when
                     sequences from EST generated from this library will be
                     masked for vector and adaptor sequences, an additional
                     masking step will have to be included to mask this RNA
                     oligo that is common to all clones (sequence
                     CGACTGACGACGAGGACACTGACATGGACTGAAGGAGTAGAAA)."
ORIGIN
```

```
QY
124 GAGAAGTTCGCCCTGAACCCCGGCCCTGCTGGAGACCACGCAGGGCCTGAAGCAGAATCATC 183
    ||||    ||||||  |||   |||||||   |||  ||| ||| || |||| ||| ||
```

Query Match     5.4%;  Score 79.8;  DB 14;  Length 1126;
Best Local Similarity   48.2%;
Matches  287;  Conservative   0;  Mismatches 302;  Indels   6;  Gaps   2;

```
Db  763 GAGATCACCGCCAACCACCCATGAAACCTGATGAGCGATCGGCGCGGACGAGGGCCAGTC 704
        |||||  ||||||||||   ||||   ||||||| ||||||||||   || |||||||
Qy  184 CGCCAGCTGCACCCCGCCCCTGCAGACCGGCCGCGAGGAGGCTGAAGAGCCTGTTCAACACC 243
```

```
Qy  244 GTGGCCACCCTGTACTGCCTGCACGGAGAAGATCGAGGTTCCGCGACACCAAGGAGGCCTG 303
        |||||||||| |||  ||||    ||  || |||| ||||||||  |||||| ||||||||
Db  703 CTCAACATCGCTGCTCAAGCTCATCGGGCCCCAAGAAGACCATGGAGATCGGCGTCTACACC 644

Qy  304 GACAAGATCGAGCGAGGAGGACGAGAACAAGTGCCAGCAGAAGATCCAGCAGGGCGAGGCC 363
        ||||||||||| |||||||| || || ||||| |||| ||||| ||| ||| || || |||
Db  643 GGCTACTCCCTCCTTCGCCACCGGCGTCGGCCCATCCCCGACGAGCGGCACCATCTTGGCCATG 584

Qy  364 GACAAGGGCAAGGTGAGCCAGGAACTACCCCATACGTTGCAGAACCTTGCAGGGTGGGG 423
        ||||||| |||||||| |||| |||| ||||   ||||  |||||||  ||| || ||
Db  583 GACATCAACCGCGAGAACTACGAGCTGGGCGTCGCCTGCGCTGCATCGAGAAGGCCGGGTGGGG 524

Qy  424 CACCAGGCCATCAGCCCCCGCACCCTGACGCGCTGGGGTGAAGGTGATCGAGGAGAAGGCC 483
        |||| || |||| |||||| || || || |||||  ||  || ||| ||| ||||| |||
Db  523 CACAAGATCGACTTCCGCGACGAGGG---CCCGGGCTGCCGTGCTGGATCGGCGCTGCTGGAG 467

Qy  484 TTCAGCCCCGGAGGTGATCCCCATGTTCACCGCCCCTGAGCGAGGGGCGCCACCCCCAGGAC 543
        ||| || ||| || ||| ||||| || ||||    ||||||||  ||| || ||| |||||
Db  466 GACGAGCGGCCAACCACGGCCACTTCGACTTCGTTCTTGCGTGGAACGCGCCAAGGACGAACTAC 407

Qy  544 CTGAACACGATGTTGAACACCGTGGGCGGCGCCACCAGGCGCCATGCAGATGCTGAAGGAC 603
        |||| |||||||| ||||||||| ||||| |||||   ||||| ||| || |||| || |
Db  406 CTCAACTACCACGACGCGCCTCATGAAGCTGGTCAAAGCTCGGCGGCGGCTTCTCGGTTACGAC 347

Qy  604 ACCATCAACGAGAGGAGGCCCGCGAGTGGGGACCCGGCTGTGCACCGCGGGCCCCATC 663
        || || |||||| ||||| || || ||||| |||| || || ||||||||| |||||
Db  346 ---AACACGCTCTGGAACGGCTCGTCCCCGCGCCGAGCCCCCCAAGCGGCCAAGTAC 290

Qy  664 GCCCCCGGCCAGATGCCGCAGCCCCGGCGCAGGGACATCGCCGGCGCACCACCAGCA 718
        ||||| ||| || ||  || |||| ||  ||||||| || ||  || ||||||  |
Db  289 ATCCCGTCCTTACCGCCGGACTTCGTCCTCGACCTCAACAGCCCCCTTCGGCCCCGACCAGGGC 230

Qy  229 GTCCAGGATCTGCCAGCTCCCCGGTCGGGACGGGCATCACCCCTCTGCGCGGCGGGCCA 175
        |||||| || | ||| ||  || |||| || |||| ||| ||| | |||||||| |||
Db  229 GTCCAGGATCTGCCAGCTCCCCGGTCGGGACGGGCATCACCCCTCTGCGCGGCGGGCCA 175
```

RESULT 6

LOCUS        CB662794         821 bp    mRNA    linear    EST 09-APR-2003
DEFINITION   OSJNBd07C09.f OSJNEd Oryza sativa (japonica cultivar-group) cDNA
             clone OSJNBd07C09 5', mRNA sequence.
ACCESSION    CB662794
VERSION      CB662794.1  GI:29666519

KEYWORDS     EST.
SOURCE       Oryza sativa (japonica cultivar-group)
  ORGANISM   Oryza sativa (japonica cultivar-group)
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE    1  (bases 1 to 821)
  AUTHORS    Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,
             Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE      Large-scale identification of ESTs involved in the interaction
             between rice and Magnaporthe grisea
  JOURNAL    Unpublished (2003)
COMMENT      Contact: Rod Wing
             Arizona Genomics Institute
             University of Arizona
             Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
             85721-0088, USA
             Tel: 520 626 3967
             Fax: 520 621 9288
             Email: http://genome.arizona.edu
             PCR Primers
             FORWARD: gta aaa cga cgg cca gtg
             BACKWARD: gga aac agc tat gac cat g
             Plate: 07  row: C  column: 09
             Seq primer: gta aaa cga cgg cca gtg.
FEATURES             Location/Qualifiers
     source          1..821
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNBd07C09"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNBd"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; 24 hrs after innoculation with Rice Blast (C9240-1)"

ORIGIN

Best Local Similarity  44.4%;
Query Match
Matches 312; Conservative  0;  Pred. No. 2.3;  DB 14;  Length 821;
                                 Mismatches 390;  Indels  0;  Gaps  0;  Score 78;

Qy    204  GCAGACCGGCAGCGAGGAGCTGAAGAGCCCTGTTCAACACCGTTGGCCACCCTGTACTCGT  263
           |||||   ||||   |||| |||| |||||||||  ||||| ||||||||||| ||| |||
Db     88  GCAGATAGAAATGGCGGCCGGAGACGTTCCTCTTCACCTCCGAGTCCGTGAACGAGGGTCA  147

Qy    264  GCACGAGAAGATCGAGGTCTCCGCGACACCAAGGAGGCCCTGGACAAAGATCGAGGAGGAGCA  323
Db    148  CCCGGACACAAGCTGTTCGGACCCAGGTGTCGACGCGGTGCCTGGACGCGGTGCCCAGGA  207

Qy    324  GAACAAGTGCCAGCAGAAGATCCAGCAGGCCCGACGAAGGGCCGAAGGTGAGCCA  383
Db    268  CGAGATCACCGACCAAGGCCACACATGTTCGGACTACGACATGTTCGGCTAGGCCCCAA  327

Qy    384  GAACTACCCCATCGTTGCAGAACCTGCAGGGACGACCATCAACGAGGAGCCGC  443
Db    384  GAACTACCCCATCGTTGCAGAACCTGCAGGGACGACCATCAACGAGGAGCCGC  443

Qy    504  CATGTTCACCGCCCCTGAGCGAGGGGCGCCACCCCCAGGAGACCTGAACACGATGTTGAACAC  563
           |||
Db    504  CATGTTCACCGCCCCTGAGCGAGGGGCGCCACCCCCAGGAGACCTGAACACGATGTTGAACAC  563

Qy    564  CGTGGGCGGCCACCAGGCCCGCCACATCGCCCGACAATCGCC  623
Db    564  CGTGGGCGGCCACCAGGCCCGCCACATCGCCCGACAATCGCC  623

Qy    624  CGAGTGGGGACCGGCGTGCACCCCGTGCACCCGCGGCCGCCCCCCGGCGCCAGATGCGCGA  683
Db    624  CGAGTGGGGACCGGCGTGCACCCCGTGCACCCGCGGCCGCCCCCCGGCGCCAGATGCGCGA  683

---

LOCUS        CB662794         821 bp    mRNA    linear    EST 09-APR-2003
DEFINITION   OSJNBd07C09.f OSJNEd Oryza sativa (japonica cultivar-group) cDNA
             clone OSJNBd07C09 5', mRNA sequence.
ACCESSION    CB662794
VERSION      CB662794.1  GI:29666519

ORIGIN

Best Local Similarity  44.4%;   Score 78;  DB 14;  Length 807;
Query Match
Matches 312; Conservative  0;  Pred. No. 2.3;  Mismatches 390;  Indels  0;  Gaps  0;

Qy    204  GCAGACCGGCAGCGAGGAGCTGAAGAGCCCTGTTCAACACCGTTGGCCACCCTGTACTCGT  263
           |||||   |||   |||| |||| |||||||||  ||||| ||||||||||| ||| |||
Db     90  GCAGATAGAAATGGCGGCCGGAGACGTTCCTCTTCACCTCCGAGTCCGTGAACGAGGGTCA  149

Qy    264  GCACGAGAAGATCGAGGTCTCCGCGACACCAAGGAGGCCCTGGACAAAGATCGAGGAGGAGCA  323
Db    150  CCCGGACACAAGCTGTTCGGACCCAGGTGTCGACGCGGTGCCTGGACGCGGTGCCCAGGA  209

Qy    324  GAACAAGTGCCAGCAGAAGATCCAGCAGGCCCGACGAAGGGCCGAAGGTGAGCCA  383
Db    330  CATCGGCCTTCGTGTCCGACGAGGTCGGGCCTCGACCGCTGCAAGGTGCTCGTCAA  389

Qy    384  GAACTACCCCATCGTTGCAGAACCTGCAGGGACGACCATCAACGAGGAGCCGC  443
Db    270  CGAGATCACCGACCAAGGCCACACGTCGACTACGACATGTTCGGCTAGGCCCCAA  329

Qy    444  CACCCTGAACGCCCTGGGGTGAAGTTGATCGAGGAGGAAGGCCTTCAGCCCCGAGGTGATCCC  503
Db    444  CACCCTGAACGCCCTGGGGTGAAGTTGATCGAGGAGGAAGGCCTTCAGCCCCGAGGTGATCCC  503

Qy    504  CATGTTCACCGCCCCTGAGCGAGGGGCGCCACCCCCAGGAGACCTGAACACGATGTTGAACAC  563
Db    504  CATGTTCACCGCCCCTGAGCGAGGGGCGCCACCCCCAGGAGACCTGAACACGATGTTGAACAC  563

Qy    564  CGTGGGCGGCCACCAGGCCCGCCACATGCAGATGCTGAAGGACACCATCAACGAGGAGCCGC  623
Db    450  CGAGGAGATCGGCGCCGGGGCCGACAGGGCCGCCACATGTTCGGCTACGCCTGCCACGACCCC  509

Qy    624  CGAGTGGGGACCGGCGTGCACCCCGTGCACCCGCGGCCGCCCCCCGGCGCCAGATGCGCGA  683
Db    510  CGAGCTGATGCCCCTCAGCGAGTCTGCCGCCAAGTCTCCGGCGGCGCCCCTCACCGAGGT  569

Qy    684  GCCCCGCCGGCGACAATCGCGCGCACCAGCACCAGGCAGGATCGCCTGGAT  743
Db    570  CCGCAAGAACGGCACCTGCCGGTCAGGCAAGGTCACCGTTGA  629

Qy    744  GACCCAGCCAACCCCCCATCCCGTGGGCGGACATCTACAAGCGGTGGATCATCCTGGGGCCT  803
Db    630  GTACCTCAAGCGACGCGGCGCCATGGTCCCCGGTCCACACCGTCCATCTCCAC  689

Qy    804  GAACAAGATCGTGCGGGGATGTACAGCGCCCCGTGAGCCGTGAGTCGTCGGCAGAGAGCGG  863
Db    690  CCAGCACGACGAGAGCCGTCACCAAGCGACGACCTCAAGGACGCACGTCAT  749

Qy    864  GGAGCGCCTTCCGCGGACTACGTGGACCCGCTTCTTCAAGACCCT  905
Db    750  CAAGCCGCGTCATCCCCGACTAGCTACTTCGACGAGAAGACCCAT  791

RESULT 7

```
Qy    204 GCAGAGACCGGCAGCGGAGGAGGGACCTGTTCAACACCGGTGGGCCACCCCTGTACTGCGGT 263
          ||||| ||| || |||| | || | ||| |||| | |  | ||| |  |||  | || |||
Db    204 GCAGAGACCGGCAGCGGAGGAGGGACCTGTTCAACACCGGTGGGCCACCCCTGTACTGCGGT 263
```

ORIGIN

Matches 312; Conservative 0; Mismatches 390; Indels 0; Gaps 0;
Best Local Similarity 44.4%; Pred. No. 2.3; DB 14; Length 824;
Query Match 5.3%; Score 78;

FEATURES             Location/Qualifiers
     source          1..824
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNBd12P19"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNBd"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; 24 hrs after innoculation with Rice Blast (C9240-1)"

Fax: 520 621 9288
Tel: 520 626 3967
85721-0088, USA
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
University of Arizona
Arizona Genomics Institute
Contact: Rod Wing
COMMENT    Unpublished (2003)
JOURNAL
           between rice and Magnaporthe grisea
TITLE      Large-scale identification of ESTs involved in the interaction
           Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
AUTHORS    Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
REFERENCE  1  (bases 1 to 824)
           Ehrhartoideae; Oryzae; Oryza.
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
ORGANISM   Oryza sativa (japonica cultivar-group)
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
SOURCE     Oryza sativa (japonica cultivar-group)
KEYWORDS   EST.
VERSION    CB666192.1  GI:29669917
ACCESSION  CB666192
DEFINITION clone OSJNBd12P19 5', mRNA sequence.
LOCUS      CB666192      824 bp    mRNA    linear   EST 09-APR-2003
```

Sequence alignment block (partial, RESULT 7):

```
Db    508 CGAGCTGATGCCTCAGCCACGTCGGCGGCCCCGCCCAAGCTTCGGCGGCGGCCCTCACCGAGT 567
Qy    684 GCCCCGCGCGGCAGCGACCATGCCCGGCACCACCCAGCGCCTGCAGCGAGCGAGATCGGCCGCCTGGAT 743
Db    568 CCGGACAGAACGGCACCTTGCGGCTGGCTCAGGGCCGACACCCTGCAGGAGCAGATGCCGGTTGA 627
Qy    744 GACCAGCAACCCCCCATCCCGTGGGGCGGGGATCATCCTGGGCCT 803
Db    628 GTACCTCAACGACGGCCCGCCCATGGCCGGCCGATGGTCCCCGGTCGGGCGTCCTCATCCTCAC 687
Qy    628 GAACCAGGACCATCAGACGGCATCAAGACGTCCCAA 683
Db    748 CAAGGCCGGTCATCCCCGACAAGTACCTCGACGAGGAGGACCAT 789
Qy    688 CCAGCACGAGAGACCGTCACCAACGACGAGGCACGTCAT 747
Qy    804 GAACACAGATCGTGCGGATGTACAGCCCGGCATCAAGACGTCAGGGCCCCCAA 863
Qy    864 GGAGCCTTCGCGGACTAGGTCGGACCACCGGTTCCTCTTCAAGACCCT 905
```

Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 07   Row: H   Column: 03
Seq primer: gta aaa cga cgg cca gtg.

FEATURES
    source
    1. .851
    /organism="Oryza sativa (japonica cultivar-group)"
    /mol_type="mRNA"
    /cultivar="Nipponbare"
    /db_xref="taxon:39947"
    /clone="OSJNRb07H03"
    /tissue_type="leaf"
    /dev_stage="3 week"
    /lab_host="DH10B"
    /clone_lib="OSJNEb"
    /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
    XhoI; 24 hrs after innoculation with Rice Blast (Che
    86061)"

ORIGIN

Query Match     5.3%;     Score 78;     DB 14;     Length 851;
Best Local Similarity   44.4%;     Pred. No. 2.4;
Matches  312;     Conservative   0;     Mismatches  390;     Indels   0;     Gaps   0;

RESULT 9
CB673355
CB673355
LOCUS           CB673355                807 bp    mRNA    linear    EST 09-APR-2003
DEFINITION      OSJNEe07N10.f OSJNEe Oryza sativa (japonica cultivar-group) cDNA
                clone OSJNEe07N10 5', mRNA sequence.
ACCESSION       CB673355
VERSION         CB673355.1    GI:29677080
KEYWORDS        EST.
SOURCE          Oryza sativa (japonica cultivar-group)
ORGANISM        Oryza sativa (japonica cultivar-group)
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                Ehrhartoideae; Oryzeae; Oryza.
REFERENCE       1  (bases 1 to 807)
AUTHORS         Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,
                Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE           Large-scale identification of ESTs involved in the interaction
                between rice and Magnaporthe grisea
JOURNAL         Unpublished (2003)
COMMENT         Contact: Rod Wing
                Arizona Genomics Institute
                University of Arizona
                Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
                85721-0088, USA
                Tel: 520 626 3967
                Fax: 520 621 9288
                Email: http://genome.arizona.edu
                PCR Primers
                FORWARD: gta aaa cga cgg cca gtg
                BACKWARD: gga aac agc tat gac cat g
                Plate: 07   Row: N   Column: 10
                Seq primer: gta aaa cga cgg cca gtg.

FEATURES
    source
    1. .807
    /organism="Oryza sativa (japonica cultivar-group)"
    /mol_type="mRNA"
    /cultivar="Nipponbare"
    /db_xref="taxon:39947"
    /clone="OSJNEe07N10"
    /tissue_type="leaf"
    /dev_stage="3 week"
    /lab_host="DH10B"
    /clone_lib="OSJNEe"
    /note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
    XhoI; 24 hrs after innoculation with Rice Blast (70-15)"

ORIGIN

Query Match     5.2%;     Score 76.6;     DB 14;     Length 807;
Best Local Similarity   44.3%;     Pred. No. 3.5;
Matches  310;     Conservative   0;     Mismatches  389;     Indels   0;     Gaps   0;

RESULT 10

LOCUS       CK163513                1138 bp    mRNA    linear      EST 05-DEC-2003
DEFINITION  PGAS016142 Triticum aestivum PGAS: Library 4 Gate 8 Triticum
            aestivum cDNA, mRNA sequence.
ACCESSION   CK163513
VERSION     CK163513.1  GI:38993817
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 1138)
  AUTHORS   Allard,P., Crosby,W.L., Danyluk,J., Budes,P., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
            Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE     Functional Genomics of Abiotic Stress in Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas_est@cs.usask.ca
            This sequence is the direct result of the Base calling software
            Phred (default parameters). It is the raw base calls. To aid in the
            identification of the high quality insert the software Lucy
            (default parameters) has been run on this sequence. Lucy identified
            the region (8,827).
            Plate: L48008  row: P  column: 04.
FEATURES             Location/Qualifiers
     source          1..1138
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /clone_lib="Triticum aestivum PGAS: Library 4 Gate 8"
                     /db_xref="taxon:4565"
                     /note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Aerial
                     parts (crown and leaf) of wheat cultivar Norstar during
                     dehydration stress. 8 mRNA populations were combined
                     before constructing the library. The first four come from

ORIGIN
      Query Match         5.2%;  Score 76.6;  DB 14;  Length 1138;
      Best Local Similarity   49.0%;  Pred. No.3.7;
      Matches 291;  Conservative   0;  Mismatches 295;  Indels   8;  Gaps   3;

Qy    124 GAGAGAGTTCGGCCTCGAGAACCCGGGCCTTGTCGGAGRACGACGACGACGACGGTTGCAAGCCAAGATCATC 183
          ||||| ||  |||| |||  |||||| || ||  || | |  | |||| || |||||| ||  ||||
Db    232 GAGATCAGCGCCTACCACCCATGGAAACCTCATGACGACGTCGGCCGACGAGGGCCAGTTC 291

Qy    184 CGCCAGCTGCACCCCGCCCTTGCAGACCGGACGAGCGAGGAGCTGAAGAGCCTGTTCAACAC 243
          || |||  |  || ||| ||| | | || | ||| || ||| || |  | ||| ||||  ||
Db    292 CTCAACATGCTGCTCAAGCTCATCGGCGCCAAGAAGACCATGGAGGTGCGTCTACACC 351

Qy    244 GTGGCCACCCTGTACTGCCGTGCACGAGAGATGGTTCCGGACACCAGGAGGCCCTG 303
          |||| |||| || | |||  || |||| ||| ||  ||||||| |||||   |||
Db    352 GGCTACTCCCTGCTCGCCACCGGCGTCCGCCATCCCGACGACGCACCATCTTGGGCATG 411

Qy    304 GACAAGATCGAGCGAGGAGGAGAACAAGTGCCAGCAGAGATCCAGCCGGCGAGGCGCC 363
          |||| || | |||||||| | |||| || ||  |||||| || ||  |||||||||||
Db    412 GACATCAACCGGCGAGAAACTACGAGCTGGGCCGTGCATCGAGAAGGCGCGGGGTGGGG 471

Qy    364 GACAAGGGCAAGGTGAGCCAGAAACTACCCCATGTGCAGAACCTGCAGGGCCAGATNGTG 423
          |||||||| || | |||||||| || || |||| ||||| ||| || || ||||  | |
Db    472 CACACAGATCGACTTCCGCGAG---GGCCCGGGCGTGCCCGTTCCTGGACCGCGGCTTGGAG 528

Qy    424 CACCAGGCCATCAGCCCCCGCCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCC 483
          || |||  ||| |||| ||||||| | |||||||| || |||||| || |||||| ||||
Db    529 GACGAGGCCAACCACGGACCTTCGACTTCGTTCTTCGTGGACGCCCGACCAAGGACAACTAC 588

Qy    484 TTCAGCCCCCAGGGTGGATCCCCATGTTCACCGGCCCCTTGAGCGAGGCGCCACCCCCCAGGAC 543
          |||  |||||| || || |||||||| |  ||| | ||| ||||| ||| ||    ||||| |||
Db    589 CTCAACTACCACAGCAGCGCGCTCATGAAGCGTGCTCAAGGTCGGCGGCGCTCTCCGGCGTACGAC 648

Qy    544 CTGAACACGAGTGAACACCGTGGGGCGGCGCCACCAGGCCGCCGCATGCAGAGATNGCTGAAGGAC 603
          |||||||||| ||||||||| | || ||| ||||| ||| | || |||| ||| ||||||||||
Db    649 ---AACACCCTCTGAACGGCTCCGTCGTCCTCCCCCGCAGCCCCCCATCGCCAAGTAC 705

Qy    604 ACCATCAACAGGAGGAGCCCGCGGCTCGACGTGGGACCGC--GTGCACCCCGTGCACCGCGGCCCA 661
          |||| || | |||| | ||| || || | |||| |||||  || ||||| ||||| ||||| |||
Db    706 ATCCGCTACTACGCGACTTCGTCCTCGAGCTCAACAGGCCCTCGCCGTCCCCGCGNGACCAGCCCG 765

Qy    662 TCGCCCCCGGCCAGGATNCGGCGCGAGCCCCGCGCAGCGACCATCGCCGNCGACCCACCGACCA 715
          |||||||||||||| | | |||| |||||| | ||||| |||| ||||  ||||||  ||||
Db    766 TCGAGATCTCCAGCTCCCCCGTGCGGCGACGGGCCATCACCCCTTCGCCCGCGCGCCA 819

──────────────────────────────────────────────────────────

RESULT 11

LOCUS       CB681973                 812 bp    mRNA    linear      EST 09-APR-2003
DEFINITION  OSJNEf08M09.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEf08M09 5', mRNA sequence.
ACCESSION   CB681973
VERSION     CB681973.1  GI:29685698
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 812)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

ORIGIN

Qy    504 CATGTTCACCGGCCCCTGAGCGAGGAGGGCGCCACCCCCCAGGACATGTTGAACAC 563
          |||||    ||||||||||||||| |||  |  |  |||||| |||| |||||||
Db    409 CATCGAGCAGCAGTCGCCCGACATCGGCGCGGGGGGGTGCAGCGGCCGCACTCACCAAGCGCC 468

Qy    564 CGTGGGCGCGGCCACCAGGCCGCGCATGCAGATGCTGAAGAGACCATCAACGAGGAGGCCGC 623
          ||||||| ||  | |||||| | || ||| ||| | |||  | |||| || |||||||  |
Db    469 CCGAGGAGATCGGGCCGCCGGGAACCAGGGCCCACATGTTCGGCTACGCCGCACGGACACCC 528

Qy    624 CCGAGTGGGACCCGGTGCACCCCGTGCACCGCGGCCCCATCGCCCCGGCGATGCGCGGA 683
          | ||||| |||| | ||||||||||||||||||||||||||||||||||||||||||||
Db    529 CGAGCTGATGCCCTCAGCCACCGTCACCGGGCCCCCTCAAGGTCGGCGGCCA 588

Qy    684 GCCCCGCGCGGCGGCGGACATGCCGCCGCCACACCCCTGCAGCGGAGAGATCCCTGGAT 743
          ||||||||||||| ||||||||||||||||||||| |  ||  ||||  |||  ||||
Db    589 CCGCAAGAACGGCGACCTTGGCCTCAGGCCGACGCAGAGGGTCACCGTTGCA 648

Qy    744 GACCAGCAACCCCCCATCCCGTGGGCGGCCGAACATCCCGGTGGATCATCCTGGGGCCT 803
          ||||||||| |||| || |||||||||||| |||||| |||| ||||| |||||| |||
Db    649 GTACCTCAACGAGACGACGGGCGGCCGTGCCGGGTCCCACACCGTCCTCTCAATCTCCAC 708

Qy    804 GAACAAGATCGGTGCGGAGNTAACAGCCCCCGTGAGCATCAACCAGGGGCCCCCCAA 863
          ||||||||||| ||||| |||  |||| || | ||| | |||  |||| || |||
Db    709 CCAGCACGAGACCGTCACCAACGACGAGATCGGCCCCGACCTCAAGGACGCACGTCAT 768

Qy    864 GGAGCCCTTCCGCGGACTACGTGGAACCGCTTCTTCAAGAC 902
          ||||| || |  |||| || ||||  || || ||   ||||
Db    769 CAAGCCCGTCATCCCCGACGAAGTNACCTGCGACGGAAGAC 807

TITLE      Large-scale identification of ESTs involved in the interaction
           between rice and Magnaporthe grisea
JOURNAL    Unpublished (2003)
COMMENT    Arizona Genomics Institute
           University of Arizona
           Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
           85721-0088, USA
           Tel: 520 626 3967
           Fax: 520 621 9288
           Email: http://genome.arizona.edu
           PCR Primers
           FORWARD: gta aaa cga cgg cca gtg
           BACKWARD: gga aac agc tat gac cat g
           Plate: 08 row: M column: 09
           Seq primer: gta aaa cga cgg cca gtg.

FEATURES   Location/Qualifiers
   source  1..812
           /organism="Oryza sativa (japonica cultivar-group)"
           /mol_type="mRNA"
           /cultivar="Nipponbare"
           /db_xref="taxon:39947"
           /clone="OSJNBf08M09"
           /tissue_type="Leaf"
           /dev_stage="3 week"
           /clone_lib="OSJNBf"
           /lab_host="DH10B"
           /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
           XhoI; Uninfected Control"

ORIGIN

Query Match    5.2%;  Score 76.4;  DB 14;  Length 812;
Best Local Similarity   44.3%;  Pred. No. 3.7;
Matches 311;  Conservative    0;  Mismatches 391;  Indels    0;  Gaps    0;

Qy    204 GCAGACCCGGCAGCGGAGGAGCTGAAGAGCCTGTTCAACACCGTGGCCACCCTGTACTGCGT 263
          ||||  ||| ||||| | |||||||  |||| || | ||||| ||||  ||||||||||||
Db    105 GCAGATAGAAATGGCGCGGCGGACGTTCCTCTCTTCACCTCCGGAGTCCGTGAACGAGGGTCA 164

Qy    264 GCACGAGAAGATCGAGGTCCGCGACACCAAGAGGCCCTGGACAAGATCGAGGAGGAGGAGCA 323
          ||| ||| ||||||||  |||  ||  ||||| || ||||||| | ||||||| |||| |||
Db    165 CCCGGACAAGCTGTCGACCAGTTGTCGGACGCGCGGTGCTCGACGTGCCTCGCCCCGCCAGGA 224

Qy    324 GAACAAGTGCCCAGCAGAAGATCCAGCACGAGGCCGGAGGCCGCCAAGGTGAGCCA 383
          |||  |||||| | | ||  | | ||| |||| | |||  ||| ||||| |||
Db    225 CCCCGACGAGCAAGGTGCGTGCCACGACGTGCCACCAACACGATGTGATGCGTTCGG 284

Qy    384 GAACTACCCCATCGTGCAGAACCTGCAGGCCCAGATAGGTGCACCAGGGCCATCAGCGCCCCG 443
          || |  ||||  ||||||  ||||||| |||| ||| |||| || |||||||| |||||||
Db    285 CGAGATCACCACCAAGGCCACCGTCGACTACGAGAAGAATCGTCCGCGACCACTGCCGCGG 344

Qy    444 CACCCTGAACCGCCTGGCCTGAAGGTGATCGAGGAGGAAGGCCTTCAGCCCCGAGGTGATCCC 503
          ||||| | |  | | || |||| |||||| | ||||| | |||||| |||  | | |||
Db    345 CATCGCGCTTCGTGTCCGACGAGCTGCCTCGGCCGCTGCAAGGTGCTCGTCGTCAA 404

Qy    504 CATGTTCACCCGCCTGAGCGGAGGCCGGCCCACCCCCGAGACCTGAACACGATGTTGAACAC 563
          || |  ||  |||| |||  |   ||| | |||   |||  | | || |||| ||||
Db    405 CATCGAGCAGCAGTCGCCCGACATCGGCGCAGGGGTGCGCCAGGCCCACTTCACCGGATGCC 464

Qy    564 CGTGGGGGCCGCCACCAGGCCCATGCAGATGCTGAACGACAGCACCATCAAGCGAGGGCCGC 623
          | ||||||| | ||| | |||| ||| || |  |||  |||| ||| | |||||||||
Db    465 CGAGGAGATCGGCGGCGGGCCGCCCACAGTGCCGGCTACGCGACCAACGAGCGAGACCC 524

Qy    624 CGAGTCGGACCCGTGCACCCCGTGCAGCGCGCCCATGCGCCCCGGCCAGATGCGCGGA 683
          || |||||  ||||| || ||||| ||| || ||||||   ||||  ||| ||| |||
Db    525 CGAGCTGAATGCCCTCAGCAGCGTCGGCGCCCGCCACCAAGCTCGGCGGCCAGGTCGAGGT 584

Qy    684 GCCCCGCCGGCAGCGACACGGTCGCCGGGCACACACAGGAGGAGCAGATGCCCTGGAT 743
          ||||| ||  | | |||| | |||| || |   ||| | ||| | ||| || |||
Db    585 CCGGCAAGAACGGCAACTTGCGCTCGGCAAGAGACCCAGGGTCACCGGTTGA 644

Qy    744 GACCAGCAACACCCCCCCCATCCCCGTGGGCCGGCGCAATTCACGAGCGGTGGATCATCCTGGCGCCT 803
          |  ||  | |||| | ||||  |  |  ||||  ||||| | |  ||||||  | |||  |
Db    645 GTACCTCAACGACGCGGCCGCCATCGGTCCCGGTTCGGGTGTCCGGCGTCGTCTTCATCTCCAC 704

Qy    804 GAACAAGATCGTGCGGATGTGACGGCATCCCGGGCGCCGGCATCCTGGACATCAAGCAGCAGGGCCCGGA 863
Db    705 CCAGCACGAGCGACCGTCACCAACGACGAGCACCCGTGAGCATGCTGGGCACAAGCAGCAGGGCCCGGA

Db    765 CAAGCCGGTCATTCCCGACAAGTACCTCGACGAGGAGACCAT 806

─────────────────────────────────────────────────────────

RESULT 12
LOCUS       CB684834                  832 bp    mRNA    linear    EST 09-APR-2003
DEFINITION  OSJNBf14L24.f OSJNBf Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNBf14L24 5', mRNA sequence.
ACCESSION   CB684834
VERSION     CB684834.1  GI:29688559
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 832)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR Primers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 14 row: L column: 24
            Seq primer: gta aaa cga cgg cca gtg.

FEATURES    Location/Qualifiers
   source   1..832
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /cultivar="Nipponbare"
            /db_xref="taxon:39947"
            /clone="OSJNBf14L24"
            /tissue_type="Leaf"
            /dev_stage="3 week"
            /clone_lib="OSJNBf"
            /lab_host="DH10B"
            /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
            XhoI; Uninfected Control"

ORIGIN

Query Match    5.2%;  Score 76.4;  DB 14;  Length 832;
Best Local Similarity   44.3%;  Pred. No. 3.7;
Matches 311;  Conservative    0;  Mismatches 391;  Indels    0;  Gaps    0;

Qy    204 GCAGACCCGGCAGCGGAGGAGCTGAAGAGCCTGTTCAACACCGTGGCCACCCTGTACTGCGT 263
          ||||  ||| ||||| | |||||||  |||| || | ||||| ||||  ||||||||||||
Db     97 GCAGATAGAAATGGCGCGGCGGACGTTCCTCTCTTCACCTCCGGAGTCCGTGAACGAGGGTCA 156

Qy    264 GCACGAGAAGATCGAGGTCGCGACACCAAGAGGCCCTGGACAAGATCGAGGAGGAGGAGCA 323
          ||| ||| |||||||   |||  ||  ||||| || ||||||| | ||||||| |||| |||
Db    157 CCCGGACAAGCTGTCGACCAGTTGTCGGACGCGCGGTGCTCGACGTGCCTCGCCCCGCCAGGA 216

Qy    324 GAACAAGTGCCAGCAGAAGATCCAGCACGAGGCCGGAGGCCGCCAAGGTGAGCCA 383
          |||  ||||| | | ||  | | ||| |||| | |||  ||| ||||| |||
Db    217 CCCGGACAGCAAGGTGCGTGCCACGACGTGCCACCAACACGATGTGATGCGTTCGG 276

RESULT 13

LOCUS       BF259495      759 bp    mRNA    linear    EST 22-OCT-2001
DEFINITION  HVSMBf0019H19f Hordeum vulgare seedling root EST library HVcDNA0007
            (Etiolated and unstressed) Hordeum vulgare subsp. vulgare cDNA
            clone HVSMBf0019H19f, mRNA sequence.
ACCESSION   BF259495
VERSION     BF259495.2  GI:13120022
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
ORGANISM    Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 759)
AUTHORS     Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D.,
            Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W.,
            Fenton,R.D., Oates,R. and Main,D.
TITLE       Development of a genetically and physically anchored EST resource
            for barley genomics: Morex unstressed seedling root cDNA library
JOURNAL     Unpublished (2001)
COMMENT     On Nov 16, 2000 this sequence version replaced gi:11188608.
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hq bases = 388
            Seq primer: AATTAACCTCACTAAAGGG
            High quality sequence stop: 758.
FEATURES             Location/Qualifiers
     source          1..759
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="Morex"
                     /sub_species="vulgare"
                     /db_xref="taxon:112509"
                     /clone="HVSMBf0019H19f"
                     /tissue_type="Seedling root"
                     /lab_host="TJC121"
                     /clone_lib="Hordeum vulgare seedling root EST library
                     HVcDNA0007 (Etiolated and unstressed)."
                     /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
                     Seeds were surface sterilized then germinated under axenic
                     conditions in the dark at room temperature on filter paper
                     with water, nystatin and cefotaxime in covered
                     crystallization dishes. Five-day old seedling roots were
                     then harvested, total RNA was prepared, poly(A) RNA was
                     purified, one primary unamplified cDNA library was made,
                     and 1 million pfu were in vivo excised to give pBluescript
                     SK(-) cDNA phagemids. These steps were performed in the TJ
                     Close laboratory at the University of California,
                     Riverside (Choi, Close, Fenton). Phagemids were plated and
                     picked at the Clemson University Genomics Institute (CUGI)
                     (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
                     preparations, DNA sequencing and sequence analysis were
                     performed at CUGI (Wing, Yu, Frisch, Henry, Simmons,
                     Oates, Rambo, Main). The sequence has been trimmed to
                     remove vector sequence and contains a minimum of 100 bases
                     of phred value 20 or above. For more details on library
                     preparation and sequence analysis see
                     http://www.genome.clemson.edu/projects/barley. To order
                     this clone see http://www.genome.clemson.edu/orders Also
                     see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                     Genetically and physically anchored EST resources for
                     barley genomics. Barley Genetics Newsletter 31:29-30.
                     (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

ORIGIN
Query Match
Best Local Similarity   47.7%;
Matches 284; Conservative   0; Mismatches 305; Indels 6; Gaps 2;
                        Score 76.2; DB 10; Length 759;
                        Pred. No.3.9;

Qy  124 GAGAAGTTCGACCCTGAACCCGGGCCTGCTCGGAGACCAGCGGAGGCTGCAAGCAGATCATC 183
            |||||  |||||||||||||||||||||          |||
Db   13 GAGATCACCCGCTACCACCCATGGAAACCTGATGCAGTCGGCGCGGACGAGGGCCAGTC   72

Qy  184 CGCCAGCTGCACCCCGCCCCTTGCAGACCGGCCGAGCGGAGGAGCTGAAGAGGCTGTTCAACACC 243
            ||||                                       ||||||||||
Db   73 CTCAACATGTCGCTCAAGCTTCATCGGCGCCAAGAAGACCATGGAGATCGGGGTTCTACACG 132

Qy  244 GTGGCCACCCTGTACTGCGTGCACGAGACGAGATCGAGGGTTCCGCGACACCGAAGGAGGCCCTG 303
            ||||||
Db  133 GGCTACTCCCTGCGTCGCACCGGCGATCCCGACGGCGACCATCTTGGCCATG 192

Qy  304 GACAAGGATCGAGGGTGAGGGAGCGAGAACAAGTGCCAGCAGAAGATCCAGCAGGCGCGAGGCGCC 363
            |||||                                           ||||
Db  193 GACAATCAACCGCGAGAACTACGAGCTGGGCGGTGCCGTGCATCGAGAAAGGCCGGGGCGTGGGCG 252

Qy  364 GACAAGGGCAAGGTGAGCCAGAACTACCCCATGGTGCAGAACCTGCAGGGCCAGATGGTG 423
            |||
Db  253 CACAAGATCGAACTTCCGGCGAGGGGCCGGCGCTC---CCCGTCCTCGACGGCCCTCCTCGAG 309

Qy  424 CACCAGGCCATCAGCCCCGCCACCCTGAACGCCTGGGGTGAAGGTGATCGAGGAGAAGGGC 483
            ||||||||||||||||||||||||||
Db  310 GACGGAGGCCAACACGGCCACCTTCGACCTTCGCTTCGTTGCGGACACGGAAGGACAAACTAC 369

Qy  484 TTCAGCGCCCGAGGTGATCCCCATGTTCACCGGGCCCTGAGCGAGGGGGCCCCCCACCCCACCAGGAC 543
            |||||||||||||||
Db  370 CTCAACTACCACAGCAGCGGCCTCATGAAGCTGTCAACGGTCGGCGGCCTTCCTCGGCGCTTACGAC 429

Qy  544 CTGAACACGGATTGTTGAACACCGTTGGGGCGGCGCCCCACCCCACCCCATGCAGATGCTTGAAGGAC 603
            ||||||||||||
Db  430 ---AACACCCTCTCTGGAACGGCTCCCTCGGTTCCCGCCGCCAGCGCCCCCATGCAGTAC 486

Qy  604 ACCATCAACGAGGAGGCCGGCGACCGGTGGACCGGCGTCGCACCCGTGCACGCGGCCCCATC 663
            |||||
Db  604 ACCATCAACGAGGAGGCCGGCGACCGGTGGACCGGCGTCGCACCCGTGCACGCGGCCCCATC 663

Qy  384 GAACTACCCCATCGTGTGCAGGAACCTGCAGGGGCCAGGATGGTGCACCAGGCCATCAGCCCCCCG 443
            |||||
Db  277 CGAGGATCACCACCAAGCCACGGTCGACTACGACAGGAGAGTCCGGCGCACACACTCGCCGCGG 336

Qy  444 CACCCTGAACGCCTGGGGTGAAGGTGATGGAGGAGGACCTTCAGCCCCCGAGGTGGGATGATCCC 503
            ||||
Db  337 CATCGACGCAGCAGTCGCCGCCGACATCCGCGCAGTGCCACGGACCTGCAAGGCGTGATCTCGTCAA 396

Qy  504 CATGTTCACCGGGCCTTCGTCGTCGACGACGTCGGCGGCTCCCTCGAGGGCTGCAAGGCGTTGAACAC 563
            |||||
Db  397 CATCGAGCAGCAGTCGGCTGCGGCCGGGTGCGCACGGCCGCACTTCACCAAGCGCC 456

Qy  564 CGTGGGGGCGGCGCCGCCCATGCAGAGTCTGAAGGACACCATCAACGACGAGGAGGGCGCGGC 623
            |||||
Db  457 CGAGGAGATCGGCCGCGGCCGCCGCGCCACCATGTTCGGGCTACGGCCACCGACGGAGACCCC 516

Qy  624 CGAGGTGGGACCCGGTGCACCCGGCGCCCCACGGCCGCGGCCCCCCGGGCCGAGATGCGCCGGA 683
            ||||
Db  517 CGAGCTGATGCCCCTCAGCCACGTCTCGGCACGCTCGGCGCCGCGCCCTCAGCCCCTCACCCGGAGGT 576

Qy  684 GCCCCCGCGGCAGCCGACATCGCGCCACCAGCACCTGCAGGAGCACGATGGCCCTGGAT 743
            |||||
Db  577 CGCAAGAACGCGCACCTGGGCTCAGGCGGGGGACACCCAGGGGTCAACCAGGTCACCGGTTGA 636

Qy  744 GACCAGCAACCCCCCCATCCCCGTGGGCGGACATCTACAAGCGGGTGGGATCATCCTGGGCCT 803
            ||||
Db  637 GTACCTCAACGACGCGCCGGCCGCCGCCATGGTCCCCGGTCGACATCAAGACCAGGGCCCCAA 696

Qy  804 GAACAAGATCGTGGGATGTACAGCCCCGTGGCCATCTGGACATCAAGACCAGGGCCCCAA 863
            ||||
Db  697 CCAGCACGAGCACAGGCCGTCACCAACGACGAGATCGCGCCACCTCAAGGACACGTCAT 756

Qy  864 GGAGGCCCTTCCGGGACTACGTGGACCGCCGTTCTTCAACGACCCT 905
            |||||
Db  757 CAACGCCCGTCATCCCCGACTACTACCTCGACGAGAAGAGACCAT 798

## RESULT 14

```
BF484304

Db  487 ATCCGCTACTACCGCGACTTCGTTNCTCGAGCCCCCGGCCCCGGGACCACGCGC 546
        ||||  |||| ||     ||  | |||       |||| |  |||||||||| ||
Qy  664 GCCCCCGGCCAGATNGGCGAGGGCGAGCCCCCGGGCGCAGGCCCACCACGGCA 718

Db  547 GTNGAGATCTGCCAGCTCCCGTCGGGACGGGCCATCACCCTCTGCCGCGCCCA 601
Qy  547 GTNGAGATCTGCCAGCTCCCGTCGGGACGGGCCATCACCCTCTGCCGCGCCCA 601
```

LOCUS       BF484304                562 bp    mRNA    linear   EST 06-DEC-2000
DEFINITION  WHE2321_P12_K23ZS Wheat pre-anthesis spike cDNA library Triticum
            aestivum cDNA clone WHE2321_P12_K23, mRNA sequence.
ACCESSION   BF484304
VERSION     BF484304.1  GI:11567605
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 562)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D.,
            Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Pre-anthesis spike cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 510559773
            Fax: 5105595818
            Email: oanderson@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1..562
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE2321_P12_K23"
                     /tissue_type="spike before anthesis"
                     /dev_stage="Adult plant"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Wheat pre-anthesis spike cDNA library"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
                     greenhouse. Whole spike with awns trimmed, white, green
                     and yellow anther were collected and total RNA, and
                     poly(A) RNA were prepared, a cDNA library was made, and
                     the cDNA clones were in vivo excised to give pBluescript
                     phagemids in the TJ Close lab (Choi, Close, Fenton) at
                     the University of California, Riverside. Plasmid DNA
                     preparations and DNA sequencing were performed in the OD
                     Anderson lab (all other authors)."

ORIGIN

```
Query Match          5.1%;   Score 75.6;  DB 10;  Length 562;
Best Local Similarity   50.4%;  Pred. No. 4.3;
Matches 270;  Conservative  0;  Mismatches 254;  Indels 12;  Gaps  3;

Qy  182 TCCGGCCAGCTGCACCCCGCCCTGCACACGGACGGGACCTGAGAGCCTGTTCAACA 241
        ||||   |||  | ||| |  ||| |  | |||| |||| || ||  ||||
Db  26  TCCTCAAACATGCTGTCATCGGGCGGCGCAAAGAGAACCATGGAGGATCGGCGGTCTACA 85

Qy  242 CCGTGGGCCACCCTGTTACTCTGGGTCACGGAGAAGAGATCAGAGGTCCGCGGCGACACCGAGGAGGCC 301
        ||| ||  |||||| | |||  ||||| ||| || || || || ||||  |||  |||||  |||||||
Db  86  CCGGGTACTCCCTGCTCGGCCACCGCGCTCGGCCATCCCGACACGGCACCATCTTCGGCCA 145
```

---

## RESULT 15

```
CB870888                    649 bp    mRNA    linear   EST 03-JUL-2003
```

```
Db  302 TGGACACAGATCGGAGGAGGGAGCAGAACACAAGTGCCAGCAGCAGAAGATCCAGCAGGGCCG 361
Qy  ...

Db  146 TGGACATCAACCCGCGAGAACTACTATCTGGGGTCGCCGGGGCTGCCGGTGCATCCAGAAGGGCGTGG 205

Db  362 CCGACAAGGGCCAAGGTGACCGAGAACTACCATCGGGCTGGGGCCACCCCCCAGG 421
Qy  206 CGACAAGATCGACTTCCGGGAGG---CCCGGCGCTGCCCGGTCCCGGGCTGCTGG 262

Qy  422 TGCACCCAGGCCCATCAGCCCCCGCACCCTGAACCGCTGGGTGAAGTGGATCGAGGAAGG 481
Db  263 AGGACGGAGGCCAACCACCGGGAACCAGCCTTCGACTTCGTTCGTGGAGCAACGGGCAAGGGACAACT 322

Qy  482 CCTTCAGCCCCCGGAGGTGATCCCCCATGTTCACCCGCCCTCAGCGACCGGGGGCCCACCCCCCAGG 541
Db  323 ACCTTCAACTACACCACGACGCCCTCATGAACCTCGTCAAGTCGCGGCGGCTTCCTCGGCTACG 382

Qy  542 ACCTGAACACGGATGTTGAACACCGTGGGGCGGCCACCAGGCGGCCATGCAGATGCTGAAGG 601
Db  383 AC--AACACCCTTCTGGAACGGCGTCCGTGTGCTGCCGCCCCGCGGACGCCCCATGCGCAAGT 439

Qy  602 ACACCATCAACGAGGAGGCGCCCAGTGGGACCCGCGTGCACGGCGGCCCCA 661
Db  440 ACATCCGCTACTA-----CCGGCACTTCGTTCGTCTTCGAGGCTCAACAGGCCCTTCGCGCCG 493

Qy  662 TCGCCCCCCCGGGGGCAGGTAGGCGGGAGGGCCGGGCGCAGGCCCATCGCCGCGGACACCACGC 717
Db  494 ACCAGCGCGTCGAGATCTCCGCCAGCTCCCGTCGGGACGGGCCATCACCCTCTGCCGCGCCCA 549
```

LOCUS       CB870888                649 bp    mRNA    linear   EST 03-JUL-2003
DEFINITION  HC15J04w CH Hordeum vulgare cDNA clone HC15J04 5-PRIME, mRNA
            sequence.
ACCESSION   CB870888
VERSION     CB870888.1  GI:30072868
KEYWORDS    EST.
SOURCE      Hordeum vulgare
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 649)
  AUTHORS   Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
            Graner,A.
  TITLE     Barley ESTs from coleoptile tissue
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Stein Nils
            Molecular Markers Group, Department Genbank
            Institute of Plant Genetics and Crop Plant Research (IPK)
            Corrensstr. 3, 06466, Gatersleben, Germany
            Tel: 039482-5522
            Fax: 039482-5595
            Email: stein@ipk-gatersleben.de
            Plate: 15  row: J  column: 4
            Insert Length: 649   Std Error: 0.00
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1..649
                     /organism="Hordeum vulgare"
                     /mol_type="mRNA"
                     /cultivar="sloop"
                     /db_xref="GABI:549289"
                     /db_xref="taxon:4513"
                     /clone="HC15J04"
                     /tissue_type="coleoptile"
                     /dev_stage="coleoptile, 3 days old"
                     /lab_host="DH10B"
                     /clone_lib="CH"
                     /note="Vector: pSPORT; Site 1: SalI (5-end of cDNA);
                     Site 2: NotI (3-end of cDNA); Due to the cloning system
                     used blue/white selection for recombinants is not 100%
                     reliable. Average insert size is 1.3 kb."

ORIGIN

```
Query Match         5.1%;    Score 75.6;  DB 14;  Length 649;
Best Local Similarity   47.7%;  Pred. No. 4.4;
Matches 284;  Conservative  0;  Mismatches 305;  Indels  6;  Gaps  2;

Qy    124 GAGAAGTTCGCCCTGAACCCGGCCTGCTCGTGGAGACCAGCGAGGGCTGCAAGCAGATCATC 183
          ||||                                                         |
Db     45 GAGATCACGGCCAACCACCACCATGAAACTGATGACGACGTCGGCGGGACCAGGGCCCAGTTC 104

Qy    184 CGCCAGCTGCACCCCGCCCTGCAGACCGGCCAGCGGAGGGAGCTGAAGAGCCTGTTCAACACC 243
          |||                                         | |||||        ||||
Db    105 CTCAAACATGCTGCTCAAGCTCATCGGCGGCCAAGAAGACCATGGAGATCGGGCGTCTACACG 164

Qy    244 GTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGGCCCTG 303
          |||||                                                        
Db    165 GGCTACTCCCTGCTCGCCACCGGGCTCGGCCATCCCGACGACGGCCACCATCTTGGCCATG 224

Qy    304 GACAAGATCGAGGAGGAGGAGCACAAGTGCCAGAAGATCCAGCAGCCCGAGGGCGCCGCC 363
          |||                                       |                  
Db    225 GACATCAACCGCGAGAACTACGAGCTGGGGCTGCCCGTGCATCCAGAAGGCCGGGCGTGGCG 284

Qy    364 GACAAGGGCCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCCAGGCCCAGAATGGTG 423
          ||||                                                          
Db    285 CACAAGANTCGACTTCCGGCGGAGGGGCGCGCTC---CCCGTCCTGGACGGCCCTTCCTCGAG 341

Qy    424 CACCAGGGCCATCAGCCCCGCACCCTGAACGCCTGGGGTGAAGGTGATCGAGGAGGAAGGCC 483
          |||||||||                                                     
Db    342 GACGAGGGCCAACCACGGCACCTTCGTCTTCGTGGACGCCGACAAAGGACAACTTAC 401

Qy    484 TTCAGCCCCGAGGTGATCCCCATNGTTCACCGCCCCTGAGCGAGGGCGCGCCCCCCAGCGAC 543
          |||||                                         | |||||        
Db    402 CTCAAACTACCACGAGCGCCTCATGAAAGCTCGTCAAAGGTCGGCGGGCCTCCTCGGCGTACGAC 461

Qy    544 CTGAAACACCGATGTTGAAACCGTNGGGCGGCCCCACCAGGCCCATCAGATNGCTGAAGGAC 603
          ||||||                                         | |||||  
Db    462 --AACGCCTCTTGGGAACGGGCTCCGTNGTGCTTCCCGCGCAGCGCCCCATNGCCGAAGTTAC 518

Qy    604 ACCATCAACGAGGAGGAGGCCGCCGAGTGGGGACCGGCCGTGCACCCGGGCCCCATC 663
          |||                                                         
Db    519 ATCCGCTTACTACCGGACTTCGTTCTCGGACCTCAACAAGGCCCTTCGCCCGGNCGACCAGCGC 578

Qy    664 GCCCCCCGCCCAGATNGCGGGAGCCCCGCGGGCAGCAGCGACATNGCCCGGGCCACCACAGCA 718
          |||||||                                                        
Db    579 GTCGAGATCTGCCAGCTCCCCGTGGGCGACGGGCATCACCTCTCTGGCGCCGCGCCA 633
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM nucleic - nucleic search, using sw model

Run on:          May 28, 2004, 08:01:50 ; Search time 234.634 Seconds
                 (without alignments)
                 7636.275 Million cell updates/sec

Title:           US-09-475-704A-2
Perfect score:   60
Sequence:        1 gacatccgcgcaggggcccaa...............tggaccgcttcttcaagacc 60

Scoring table:   IDENTITY NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        27513289 seqs, 1493109276 residues

Total number of hits satisfying chosen parameters:    55026578

Post-processing: Minimum Match 0%
Maximum DB seq length: 200000000
Minimum DB seq length: 0
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
                 1: em_estba:*
                 2: em_esthum:*
                 3: em_estin:*
                 4: em_estmu:*
                 5: em_estov:*
                 6: em_estpl:*
                 7: em_estro:*
                 8: em_htc:*
                 9: gb_est1:*
                 10: gb_est2:*
                 11: gb_htc:*
                 12: gb_est3:*
                 13: gb_est4:*
                 14: gb_est5:*
                 15: em_estfun:*
                 16: em_eston:*
                 17: em_gss_hum:*
                 18: em_gss_inv:*
                 19: em_gss_pln:*
                 20: em_gss_vrt:*
                 21: em_gss_fun:*
                 22: em_gss_nam:*
                 23: em_gss_mus:*
                 24: em_gss_rod:*
                 25: em_gss_pro:*
                 26: em_gss_phg:*
                 27: em_gss_vrl:*
                 28: gb_gss1:*
                 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 29.8 | 49.7 | 646 | 29 | CC821047 | CC821047 FBCP4B6 U |
| 2 | 29.8 | 49.7 | 686 | 14 | CA180823 | CA180823 SCACST316 |
| 3 | 29.6 | 49.3 | 505 | 28 | BH786628 | BH786628 fzmb014f0 |
| 4 | 29 | 48.3 | 1151 | 13 | BQ619287 | BQ619287 RNOSEQ5H1 |

| | | | | | |
|---|---|---|---|---|---|
| 5 | 29 | 48.3 | 1151 | 13 | BQ619288 | BQ619288 RNOSEQ6A0 |
| 6 | 29 | 48.3 | 1151 | 13 | BQ619295 | BQ619295 RNOSEQ6A0 |
| 7 | 29 | 48.3 | 1151 | 13 | BQ619376 | BQ619376 RNOSEQ7A0 |
| 8 | 29 | 48.3 | 1151 | 13 | BQ619455 | BQ619455 RNOSEQ7H0 |
| 9 | 29 | 48.3 | 1151 | 13 | BQ619551 | BQ619551 C1l1astPC |
| 10 | 28.6 | 47.7 | 709 | 13 | CA072199 | CA072199 SCCCAM100 |
| 11 | 28.6 | 47.7 | 718 | 12 | BG647389 | BG647389 BST509008 |
| 12 | 28.4 | 47.3 | 525 | 12 | BG549181 | BG549181 94707JD08 |
| 13 | 28.4 | 47.3 | 544 | 14 | CA210513 | CA210513 SCEPSB112 |
| 14 | 28.4 | 47.3 | 577 | 29 | CG250176 | CG250176 OGVCM47TV |
| 15 | 28.4 | 47.3 | 580 | 12 | BG517389 | BG517389 94706TEH |
| 16 | 28.4 | 47.3 | 592 | 29 | CG226900 | CG226900 OGWIM66TH |
| 17 | 28.4 | 47.3 | 857 | 29 | CG327425 | CG327425 OGXDB776TV |
| 18 | 28.4 | 47.3 | 744 | 29 | BG550427 | BG550427 94707JH04 |
| 19 | 28.2 | 47.0 | 408 | 14 | CB639810 | CB639810 OSJNBa12B |
| 20 | 28.2 | 47.0 | 576 | 29 | CG225792 | CG225792 OG0CF27TH |
| 21 | 28.2 | 47.0 | 576 | 29 | CG225799 | CG225799 OG0CF27TV |
| 22 | 28.2 | 47.0 | 585 | 10 | AW1470097 | AW1470097 70701IF07 |
| 23 | 28.2 | 47.0 | 597 | 29 | CC675574 | CC675574 OGJAX28TV |
| 24 | 28.2 | 47.0 | 675 | 29 | CG249691 | CG249691 OGXCO69TH |
| 25 | 28.2 | 47.0 | 819 | 29 | CG293655 | CG293655 OG0BV73TH |
| 26 | 28.2 | 47.0 | 823 | 29 | CG349088 | CG349088 OGYCL45TV |
| 27 | 28.2 | 47.0 | 829 | 29 | CC882396 | CC882396 OGLB77TV |
| 28 | 28.2 | 47.0 | 829 | 29 | CC882932 | CC882932 OGYBL67TV |
| 29 | 28.2 | 47.0 | 858 | 29 | CG313210 | CG313210 OGXDR22TV |
| 30 | 28.2 | 47.0 | 869 | 29 | CC364851 | CC364851 OG1AE85TV |
| 31 | 28.2 | 47.0 | 923 | 29 | CC707427 | CC707427 OGUB52TV |
| 32 | 47.0 | 923 | 29 | CC731054 | CC731054 OGVBI45TV |
| 33 | 28 | 46.7 | 500 | 14 | CA645640 | CA645640 wrein.pk0 |
| 34 | 28 | 46.7 | 524 | 12 | BG313886 | BG313886 WHE2066_B |
| 35 | 28 | 46.7 | 615 | 14 | CD905733 | CD905733 G468.102L |
| 36 | 28 | 46.7 | 615 | 28 | CC415368 | CC415368 PUHEP34TD |
| 37 | 28 | 46.7 | 629 | 14 | CA731056 | CA731056 wip1c.pk0 |
| 38 | 28 | 46.7 | 647 | 10 | BF473271 | BF473271 WHE0926_B |
| 39 | 28 | 46.7 | 656 | 29 | CG709393 | CG709393 1119012H1 |
| 40 | 28 | 46.7 | 658 | 29 | CG310514 | CG310514 OG4BM79TC |
| 41 | 28 | 46.7 | 710 | 14 | CD932953 | CD932953 GR45.119H |
| 42 | 28 | 46.7 | 763 | 28 | BZ813699 | BZ813699 PUGFA26TB |
| 43 | 28 | 46.7 | 777 | 29 | CG240255 | CG240255 QG5AL58TC |
| 44 | 28 | 46.7 | 784 | 29 | CG204197 | CG204197 PUFKG68TD |
| 45 | 28 | 46.7 | 807 | 29 | CC667766 | CC667766 OGWBO49TV |

ALIGNMENTS

RESULT 1
CC821047/c

LOCUS           CC821047          646 bp     DNA      linear     GSS 22-OCT-2003
DEFINITION      FBCP4B6 Uncultured human fecal virus uncultured human fecal virus
                genomic, genomic survey sequence.
ACCESSION       CC821047
VERSION         CC821047.1  GI:37806847
KEYWORDS        GSS.
SOURCE          uncultured human fecal virus
  ORGANISM      uncultured human fecal virus
                Viruses; environmental samples.
REFERENCE       1  (bases 1 to 646)
  AUTHORS       Breitbart,M., Hewson,I., Felts,B., Mahaffy,J.M., Nulton,J.,
                Salamon,P. and Rohwer,F.
  TITLE         Metagenomic analyses of an uncultured viral community from human
                feces
  JOURNAL       J. Bacteriol. 185 (20), 6220-6223 (2003)
  MEDLINE       22887472
   PUBMED       14526037
  COMMENT       Contact: Rohwer F
                Biology Dept.
                San Diego State University
                5500 Campanile Dr, San Diego, CA 92102, USA
                Tel: 6195941336
                Fax: 6195995676
                Email: forest@sunstroke.sdsu.edu
                Class: shotgun.

FEATURES
```
     source          1. .646
                     /organism="uncultured human fecal virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:233364"
                     /clone_lib="Uncultured human fecal virus"
ORIGIN
```

Qy
Db
```
Query Match     49.7%;   Score 29.8;  DB 29;  Length 646;
Best Local Similarity   70.2%;   Pred. No. 9.1e+02;
Matches 40;  Conservative 0;  Mismatches 17;  Indels 0;  Gaps 0;

Qy    1 GACATCCGCCAGGGCGCCCAAGGAGAGCCCTTCCGCGGACTACGTGGACGTGGACGTTCTTCAAG 57
        |||  || ||| ||| ||| | |||||| ||||||||||||||||| || ||  ||||||||||
Db  453 GACACCCAACACGGGTCAAGCAGCCCCTTTATCGACTACGTGCCTGCCCCTTGATCCAG 397
```

RESULT 2
CA180823
```
LOCUS       CA180823                 686 bp    mRNA      linear   EST 24-SEP-2003
DEFINITION  SCACST3160C06.g ST3 Saccharum Officinarum cDNA clone SCACST3160C06
            5', mRNA sequence.
ACCESSION   CA180823
VERSION     CA180823.1  GI:35114987
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE   1 (bases 1 to 686)
  AUTHORS   Vettore,A.L., da Silva,P.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 160  row: C  column: 06
            Seq primer: T7 Promoter Primer.
```

FEATURES
```
     source          1. .686
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCACST3160C06"
                     /clone_lib="ST3"
                     /lab_host="DH10B"
                     /note="Organ: Fourth apical stalk internodes of adult
                     plants; Vector: pSport1; Site_1: SalI; Site 2: NotI; An
                     unidirectional cDNA library generated from [fourth apical
                     stalk internodes of adult plants]. cDNA was prepared from
                     polyA+ mRNA using SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were fractionated
                     in a sepharose Cl-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details of each source of RNA and library
                     construction can be obtained at
                     http://sucet.lad.ic.unicamp.br/public"
ORIGIN
```

Qy
```
Query Match     49.7%;   Score 29.8;  DB 14;  Length 686;
Best Local Similarity   70.2%;   Pred. No. 9.3e+02;
Matches 40;  Conservative 0;  Mismatches 17;  Indels 0;  Gaps 0;

Qy    1 GACATCCCGCCAGGGCGCCCAAGGAGCCCTTCCGCGGACTACGTGGACCGCTTCTTCAAG 57
        ||| || |||| ||| |||||||| |||||||||||||||||||||||| ||||||||||
```

Db
```
     11 GCCATGGACCGGGACCTCAAGGCCCGCGTCCGGCGCCGGGCCGACTGGGAGAGCTTCCTCAAG 67
```

RESULT 3
BH786628
```
LOCUS       BH786628                 505 bp    DNA       linear   GSS 28-MAR-2002
DEFINITION  fzmb014f01d05k0 fzmb filtered library Zea mays genomic clone
            fzmb014f01d05 5', genomic survey sequence.
ACCESSION   BH786628
VERSION     BH786628.1  GI:19792775
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 505)
  AUTHORS   Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
  TITLE     GeneThresher methylation filtered genomic sequences from maize
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Bedell JA
            Orion Genomics, LLC
            4041 Forest Park Ave, St. Louis, MO 63108, USA
            Tel: 314 615 6979
            Fax: 314 615 5975
            Email: jbedell@oriongenomics.com
            Plate: fzmb014f016  row: d  column: 05
            Seq primer: SK reverse
            Class: shotgun
            High quality sequence stop: 505.
```

FEATURES
```
     source          1. .505
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="Mo17"
                     /db_xref="taxon:4577"
                     /clone="fzmb014f016d05"
                     /clone_lib="fzmb filtered library"
                     /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
                     prepared from purified nuclei was randomly sheared,
                     end-repaired, size fractionated into the 0.5 to
                     5 kb fraction, ligated into HincII-digested pBCSK(-)
                     vector and electroporated into E. coli cells."
ORIGIN
```

Qy
```
Query Match     49.3%;   Score 29.6;  DB 28;  Length 505;
Best Local Similarity   68.3%;   Pred. No. 9.7e+02;
Matches 41;  Conservative 0;  Mismatches 19;  Indels 0;  Gaps 0;

Qy    1 GACATCCGCCAGGGCGCCCAAGGAGAGCCCTTCCGCGGACTACGTGGACCGCTTCTTCAAGACC 60
        |||| | | || ||| |||||| || ||||| |||| |||| | || || ||||||||||| ||
```
Db
```
     70 GCCCTCCTCCGACGCCACGAGGAGGCTTCTTCGTGGCGGCGTCGACGACGCCTTCTTCTTGACGGCC 129
```

RESULT 4
BQ619287
```
LOCUS       BQ619287                1151 bp    mRNA      linear   EST 06-SEP-2002
DEFINITION  RNOSEQ5H12_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
            mays cDNA clone RNOSEQ5H12_SK.ab1 similar to (AF287276) chlorophyll
            a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION   BQ619287
VERSION     BQ619287.1  GI:21621281
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 1151)
  AUTHORS   Wang,H. and Bohnert,H.J.
  TITLE     Genomics of plant stress tolerance
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Mark Fredricksen
```

RESULT 5

```
LOCUS         BQ619288              1151 bp    mRNA    linear    EST 06-SEP-2002
DEFINITION    RNOSBQ6A01_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
              mays cDNA clone RNOSBQ6A01_SK.ab1 similar to (AF287276) chlorophyll
              a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION     BQ619288
VERSION       BQ619288.1  GI:21621282
KEYWORDS      EST.
SOURCE        Zea mays
  ORGANISM    Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE     1  (bases 1 to 1151)
  AUTHORS     Wang,H. and Bohnert,H.J.
  TITLE       Genomics of plant stress tolerance
  JOURNAL     Unpublished (2002)
COMMENT       Contact: Mark Fredricksen
              Department of Plant Biology
              University of Illinois
              1201 W Gregory Dr., Urbana, IL 61801, USA
              Tel: 2172655473
              Email: bohnertlab@life.uiuc.edu.
FEATURES             Location/Qualifiers
     source          1..1151
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="RNOSBQ6A01_SK.ab1"
                     /tissue_type="Roots"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed Zea mays roots cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150
                     mM NaCl"
ORIGIN
```

Query Match    48.3%;  Score 29;  DB 13;  Length 1151;
Best Local Similarity   71.7%;  Pred. No. 1.7e+03;
Matches   38;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

```
Qy          8 GCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC   60
              |||||||| |||| ||||||||||||||||||||||||||||| ||||||||||||
Db        523 GGCAGGACATCAAGAGAACCCCGTCAGGCGTCAACCAGGAGGACCCCATCTTCAAGAGC  575
```

RESULT 6

```
LOCUS         BQ619295              1151 bp    mRNA    linear    EST 06-SEP-2002
DEFINITION    RNOSBQ6A08_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
              mays cDNA clone RNOSBQ6A08_SK.ab1 similar to (AF287276) chlorophyll
              a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION     BQ619295
VERSION       BQ619295.1  GI:21621289
KEYWORDS      EST.
SOURCE        Zea mays
  ORGANISM    Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE     1  (bases 1 to 1151)
  AUTHORS     Wang,H. and Bohnert,H.J.
  TITLE       Genomics of plant stress tolerance
  JOURNAL     Unpublished (2002)
COMMENT       Contact: Mark Fredricksen
              Department of Plant Biology
              University of Illinois
              1201 W. Gregory Dr., Urbana, IL 61801, USA
              Tel: 2172655473
              Email: bohnertlab@life.uiuc.edu.
FEATURES             Location/Qualifiers
     source          1..1151
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="RNOSBQ6A08_SK.ab1"
                     /tissue_type="Roots"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed Zea mays roots cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150
                     mM NaCl"
ORIGIN
```

Query Match    48.3%;  Score 29;  DB 13;  Length 1151;
Best Local Similarity   71.7%;  Pred. No. 1.7e+03;
Matches   38;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

```
Qy          8 GCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC   60
              |||||||| |||| ||||||||||||||||||||||||||||| ||||||||||||
Db        523 GGCAGGACATCAAGAGAACCCCGTCAGGCGTCAACCAGGAGGACCCCATCTTCAAGAGC  575
```

RESULT 7

```
LOCUS         BQ619376              1151 bp    mRNA    linear    EST 06-SEP-2002
DEFINITION    RNOSBQ7A07_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
              mays cDNA clone RNOSBQ7A07_SK.ab1 similar to (AF287276) chlorophyll
              a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION     BQ619376
VERSION       BQ619376.1  GI:21621370
KEYWORDS      EST.
SOURCE        Zea mays
  ORGANISM    Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE     1  (bases 1 to 1151)
  AUTHORS     Wang,H. and Bohnert,H.J.
  TITLE       Genomics of plant stress tolerance
  JOURNAL     Unpublished (2002)
COMMENT       Contact: Mark Fredricksen
              Department of Plant Biology
              University of Illinois
              1201 W. Gregory Dr., Urbana, IL 61801, USA
              Tel: 2172655473
              Email: bohnertlab@life.uiuc.edu.
FEATURES             Location/Qualifiers
     source          1..1151
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="RNOSBQ7A07_SK.ab1"
                     /tissue_type="Roots"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed Zea mays roots cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150
                     mM NaCl"
ORIGIN
```

Query Match    48.3%;  Score 29;  DB 13;  Length 1151;
Best Local Similarity   71.7%;  Pred. No. 1.7e+03;
Matches   38;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

```
Qy          8 GCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC   60
              |||||||| |||| ||||||||||||||||||||||||||||| ||||||||||||
Db        523 GGCAGGACATCAAGAGAACCCCGTCAGGCGTCAACCAGGAGGACCCCATCTTCAAGAGC  575
```

DEFINITION   CIllastPCRs_SK.ab1 Salt stressed Zea mays leaves cDNA library Zea mays cDNA clone CIllastPCRs_SK.ab1 similar to (AF287276) chlorophyll a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION    BQ619551
VERSION      BQ619551.1  GI:21621545
KEYWORDS     EST.
SOURCE       Zea mays
  ORGANISM   Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1  (bases 1 to 1151)
  AUTHORS    Wang,H. and Bohnert,H.J.
  TITLE      Genomics of plant stress tolerance
  JOURNAL    Unpublished (2002)
  COMMENT    Contact: Mark Fredricksen
             Department of Plant Biology
             University of Illinois
             1201 W. Gregory Dr., Urbana, IL 61801, USA
             Tel: 2172655473
             Email: bohnertlab@life.uiuc.edu.
FEATURES             Location/Qualifiers
     source          1. .1151
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="CIllastPCRs_SK.ab1"
                     /tissue_type="Leaves"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed Zea mays leaves cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"
ORIGIN

Qy      8 GCCAGGGGCCCAAGGAGGCCCTTCCGCGACTAGTGGACCGCTTCTTCAAGACC 60
          ||||  ||  || ||| | ||| ||  || ||| || || ||| ||||||| |
Db    523 GGCAGGACATCAAGAACACCCGTCAGCGTCAACCAGGACCCCATCTTCAAGAGC 575

RESULT 10
LOCUS        CA072199              709 bp    mRNA    linear    EST 23-SEP-2003
DEFINITION   SCCCAM1004003.g AM1 Saccharum officinarum cDNA clone SCCCAM1004D03 5', mRNA sequence.
ACCESSION    CA072199
VERSION      CA072199.1  GI:34924350
KEYWORDS     EST.
SOURCE       Saccharum officinarum
  ORGANISM   Saccharum officinarum
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE    1  (bases 1 to 709)
  AUTHORS    Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE      The libraries that made SUCEST
  JOURNAL    Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
  COMMENT    Contact: Arruda P
             Centro de Biologia Molecular e Engenharia Genetica
             Universidade Estadual de Campinas
             Caixa Postal 6010, 13083-970, Campinas SP, Brazil
             Tel: 55 19 3788 1137
             Fax: 55 19 3788 1089
             Email: parruda@unicamp.br
             Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br
             Plate: 004  row: D  column: 03

Query Match    48.3%;  Score 29;  DB 13;  Length 1151;
Best Local Similarity   71.7%;  Pred. No. 1.7e+03;
Matches 38;  Conservative  0;  Mismatches 15;  Indels  0;  Gaps  0;

Qy      8 GCCAGGGGCCCAAGGAGGCCCTTCCGCGACTAGTGGACCGCTTCTTCAAGACC 60
          ||||  ||  || ||| | ||| ||  || ||| || || ||| ||||||| |
Db    523 GGCAGGACATCAAGAACACCCGTCAGCGTCAACCAGGACCCCATCTTCAAGAGC 575

---

FEATURES             Location/Qualifiers
     source          1. .1151
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="RNOSBQ7A07_SK.ab1"
                     /tissue_type="Roots"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed Zea mays roots cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"
ORIGIN

Qy      8 GCCAGGGGCCCAAGGAGGCCCTTCCGCGACTAGTGGACCGCTTCTTCAAGACC 60
          ||||  ||  || ||| | ||| ||  || ||| || || ||| ||||||| |
Db    523 GGCAGGACATCAAGAACACCCGTCAGCGTCAACCAGGACCCCATCTTCAAGAGC 575

RESULT 8
LOCUS        BQ619455              1151 bp    mRNA    linear    EST 06-SEP-2002
DEFINITION   RNOSBQ7H05_SK.ab1 Salt stressed Zea mays roots cDNA library Zea mays cDNA clone RNOSBQ7H05_SK.ab1 similar to (AF287276) chlorophyll a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION    BQ619455
VERSION      BQ619455.1  GI:21621449
KEYWORDS     EST.
SOURCE       Zea mays
  ORGANISM   Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1  (bases 1 to 1151)
  AUTHORS    Wang,H. and Bohnert,H.J.
  TITLE      Genomics of plant stress tolerance
  JOURNAL    Unpublished (2002)
  COMMENT    Contact: Mark Fredricksen
             Department of Plant Biology
             University of Illinois
             1201 W. Gregory Dr., Urbana, IL 61801, USA
             Tel: 2172655473
             Email: bohnertlab@life.uiuc.edu.

Query Match    48.3%;  Score 29;  DB 13;  Length 1151;
Best Local Similarity   71.7%;  Pred. No. 1.7e+03;
Matches 38;  Conservative  0;  Mismatches 15;  Indels  0;  Gaps  0;

Qy      8 GCCAGGGGCCCAAGGAGGCCCTTCCGCGACTAGTGGACCGCTTCTTCAAGACC 60
          ||||  ||  || ||| | ||| ||  || ||| || || ||| ||||||| |
Db    523 GGCAGGACATCAAGAACACCCGTCAGCGTCAACCAGGACCCCATCTTCAAGAGC 575

RESULT 9
LOCUS        BQ619551              1151 bp    mRNA    linear    EST 06-SEP-2002

FEATURES
source

Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..709
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCCAM1004D03"
/clone_lib="AM1"
/lab_host="DH10B"
/clone_lib="AM1"
/note="Organ: Apical meristem and tissues surrounding of
mature plants; Vector: pSporti; Site_1: SalI; Site_2:
NotI; An unidirectional cDNA library generated from
[Apical meristem and tissues surrounding of mature
plants]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-column and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Qy      1 GACATCCGGCCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCCGCTTCTTCAAGAC 59
            ||||| ||| ||| ||||| ||||||| |||||||||||||||||| |||||||||||||||
Db    509 GACATCTTCGCGGACGGGGAAGGCCCCCCTTCGGGGACTACCTGGCCGGGGAATCCGCCAAGCC 567

RESULT 11
BG647389
LOCUS        BG647389              718 bp    mRNA    linear    EST 24-APR-2001
DEFINITION   EST650009008 HOGA Medicago truncatula cDNA clone pHOGA-16B18 5' end,
             mRNA sequence.
ACCESSION    BG647389
VERSION      BG647389.1  GI:13782501
KEYWORDS     EST.
SOURCE       Medicago truncatula
ORGANISM     Medicago truncatula (barrel medic)
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
             Medicago.
REFERENCE    1  (bases 1 to 718)
AUTHORS      Hahn,M.G., Ojanen-Reuhs,T., Samac,D., Town,C.D., Van Aken,S.,
             Utterback,T., Cho,J. and Fraser,C.M.
TITLE        ESTs from roots of Medicago truncatula treated with
             oligogalacturonides of DP 6-20
JOURNAL      Unpublished (2001)
COMMENT      Contact: Michael G. Hahn
             Complex Carbohydrate Research Center
             University of Georgia
             220 Riverbend Road, Athens, GA 30602-4712, USA
             Tel: 706-542-4457
             Fax: 706-542-4412
             Email: hahn@ccrc.uga.edu
             G39095e TIGR sequence name: MTMB2097TK More information is
             available at: www.medicago.org
             Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES
source

Location/Qualifiers
1..718
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="pHOGA-16B18"
/tissue_type="3 day old seedling roots"
/dev_stage="24 hours after treatment in the dark at 26 C
with 0.5 mg/ml oligogalacturonides (DP 6-20) in the
presence of 100 ug/ml Gentamicin"
/lab_host="XL0LR"

ORIGIN

Query Match      47.7%;    Score 28.6;  DB 13;  Length 709;
Best Local Similarity   67.8%;    Pred. No. 1.9e+03;
Matches   40;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

---

ORIGIN

/clone_lib="HOGA"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in SOLR cells."

RESULT 12
BG549181
LOCUS        BG549181              525 bp    mRNA    linear    EST 05-APR-2001
DEFINITION   947073D08.y1 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA
             sequence.
ACCESSION    BG549181
VERSION      BG549181.1  GI:13557825
KEYWORDS     EST.
SOURCE       Zea mays
ORGANISM     Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
             clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1  (bases 1 to 525)
AUTHORS      Walbot,V.
TITLE        Maize ESTs from various cDNA libraries sequenced at Stanford
             University
JOURNAL      Unpublished (1999)
COMMENT      Contact: Walbot V.
             Department of Biological Sciences
             Stanford University
             855 California Ave, Palo Alto, CA 94304, USA
             Tel: 650 723 2227
             Fax: 650 725 8221
             Email: walbot@stanford.edu
             Plate: 947073  row: D  column: 08.
FEATURES
source

Location/Qualifiers
1..525
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/tissue_type="leaf and stem, including leaf base"
/dev_stage="2 week old seedling (3 leaves)"
/clone_lib="Xll-Blue"
/lab_host="Xll-Blue"
/note="Organ: shoot; Vector: Lambda ZAP (pBluescript SK-);
Site_1: EcoRI; Site_2: XhoI; Directionally cloned using
Stratagene's UniZap XR cDNA cloning kit with the 5' end
at the EcoRI site. The library represents 8 x 10e5
independent recombinant phage. The plants were greenhouse
grown."

ORIGIN

Qy      2 ACATCCGGCCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCCGCTTCTTCAAGAC 60
            ||||| |||  ||| ||||| |||||| |||||||||||||||||||| | ||||| |||
Db    584 ACGTCCGCCAAGACCCCAGACCCAAACGAGTCCTTCCAGAATACCTCGTCTGATTCAACAACGCC 642

Query Match      47.7%;    Score 28.6;  DB 12;  Length 718;
Best Local Similarity   67.8%;    Pred. No. 1.9e+03;
Matches   40;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

RESULT 13

Qy      7 CGGCCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCCGGCTTCTTCAAGAGC 60
            |||| |||||| ||||||   |||| ||| ||||||||||||||| | ||||||||||||||
Db    282 CGGCCAGGACATCAAGAACCCCCGCCAGCCGCTCAACAGGAGACCCCAPCTTCAAGAGC 335

Query Match      47.3%;    Score 28.4;  DB 12;  Length 525;
Best Local Similarity   70.4%;    Pred. No. 2e+03;
Matches   38;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

CA210513
```
LOCUS       CA210513                544 bp    mRNA    linear   EST 25-SEP-2003
DEFINITION  SCBFSB1128C11.g SB1 Saccharum officinarum cDNA clone SCBPSB1128C11
            5', mRNA sequence.
ACCESSION   CA210513
VERSION     CA210513.1  GI:35253896
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE   1  (bases 1 to 544)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 128  row: C  column: 11
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1..544
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCBPSB1128C11"
                     /clone_lib="SB1"
                     /lab_host="DH10B"
                     /note="Organ: Stalk Bark from adult plants; Vector:
                     pSport1; Site 1: SalI; Site 2: NotI; An unidirectional
                     cDNA library generated from [Stalk Bark from adult
                     plants]. cDNA was prepared from polyA+ mRNA using
                     SuperScript Plasmid System Kit (Invitrogen). The
                     double-strand cDNAs were fractionated in a sepharose
                     CL-2B 40cm-columns and fragments sizing between 0.8 and
                     1.5 Kb were directionally cloned into the vector. Details
                     of each source of RNA and library construction can be
                     obtained at http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Query Match    47.3%;   Score 28.4;  DB 14;  Length 544;
Best Local Similarity   70.4%;   Pred. No. 2e+03;
Matches  38;   Conservative   0;  Mismatches  16;  Indels  0;  Gaps  0;

Qy    1 GACATCCGGCCAGGGCCCCAAGGAGGCCCTTCCGCGACTAGGTGGACCGGCTTCTTC 54
        |||| ||||| | ||| |||||||||| |||||||||||||| ||| ||| |||||||
Db  254 GACTTCCGCCGCGGCGGACCTCGAGGACCTCGTCTGCGACCACCTCGGGCGGGCTTC 307
```

RESULT 14
CG250176
```
LOCUS       CG250176                577 bp    DNA    linear   GSS 25-AUG-2003
DEFINITION  OGVCM47TV ZM_0.7_1.5 KB Zea mays genomic clone ZMMBMa0507H21,
            genomic survey sequence.
ACCESSION   CG250176
VERSION     CG250176.1  GI:34152266
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..577
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0507H21"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

Query Match    47.3%;   Score 28.4;  DB 29;  Length 577;
Best Local Similarity   70.4%;   Pred. No. 2e+03;
Matches  38;   Conservative   0;  Mismatches  16;  Indels  0;  Gaps  0;

Qy    1 GACATCCGGCCAGGGCCCCAAGGAGGAGCCCTTCCGCGACTAGGTGGACCGGCTTCTTC 54
        |||| ||||| | ||| |||||||||||| |||||||||||| ||| ||| |||||||
Db  378 GACTTCCGCCGCGGCGGACCTCGAGGACCTCGTCTGCGACCACCTCGGGCGGGCTTC 431
```

RESULT 15
BG517389
```
LOCUS       BG517389                580 bp    mRNA    linear   EST 30-MAR-2001
DEFINITION  947061G10.y1 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA
            sequence.
ACCESSION   BG517389
VERSION     BG517389.1  GI:13490625
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
  JOURNAL   University
            Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 947061  row: G  column: 10.
FEATURES             Location/Qualifiers
     source          1..580
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="taxon:4577"
                     /cultivar="B73"
                     /tissue_type="leaf and stem, including leaf base"
                     /dev_stage="2 week old seedling (3 leaves)"
                     /clone="947 - 2 week shoot from Barkan lab"
                     /lab_host="Xl1-Blue"
                     /note="Organ: shoot; Vector: Lambda ZAP (pBlueScript SK-);
                     Site 1: EcoRI; Site 2: XhoI; Directionally cloned using
                     Stratagene's UniZap XR cDNA cloning kit with the 5' end
                     at the EcoRI site. The library represents 8 x 10e5
                     independent recombinant phage. The plants were greenhouse
```

ORIGIN

grown."

Query Match

Best local Similarity    47.3%;    Score 28.4;    DB 12;    Length 580;
Matches    38;    Conservative    70.4%;    Pred. No. 2e+03;
                   0;    Mismatches    16;    Indels    0;    Gaps    0;

Qy      7 CGCCAGGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60
          || |||||| |--|||||||-|   |||||||||||||| |--| ||||-||||||||
Db    508 CGGCAGGACATCAAGAACCCCGGGGCAGCCGTCAAACCAGGACCCCATCTTCAAGAGC 561

Search completed: May 28, 2004, 11:33:37
Job time : 237.634 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:            May 28, 2004, 09:24:06 ; Search time 6.49156 Seconds
                   (without alignments)
                   5129.286 Million cell updates/sec

Title:             US-09-475-704A-1
Perfect score:     60
Sequence:          1 gacatcaagcaggggcccaa..........tggaccgcttcttcaagacc 60

Scoring table:     IDENTITY NUC
                   Gapop 10.0 , Gapext 1.0

Searched:          682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:   1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
           1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
           2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
           3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
           4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
           5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
           6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 53.6 | 89.3 | 60 | 4 | US-09-475-515-20 | Sequence 20, Appl |
| 2 | 53.6 | 89.3 | 1268 | 4 | US-09-475-515-9 | Sequence 9, Appli |
| 3 | 53.6 | 89.3 | 1515 | 4 | US-09-475-515-4 | Sequence 4, Appli |
| 4 | 53.6 | 89.3 | 1853 | 4 | US-09-475-515-5 | Sequence 5, Appli |
| 5 | 53.6 | 89.3 | 1865 | 4 | US-09-475-515-78 | Sequence 78, Appl |
| 6 | 53.6 | 89.3 | 1865 | 4 | US-09-475-515-79 | Sequence 79, Appl |
| 7 | 53.6 | 89.3 | 2031 | 4 | US-09-475-515-7 | Sequence 7, Appli |
| 8 | 53.6 | 89.3 | 4319 | 4 | US-09-475-515-6 | Sequence 6, Appli |
| 9 | 53.6 | 89.3 | 4472 | 4 | US-09-475-515-75 | Sequence 75, Appl |
| 10 | 53.6 | 89.3 | 4608 | 4 | US-09-475-515-76 | Sequence 76, Appl |
| 11 | 53.6 | 89.3 | 4689 | 4 | US-09-475-515-74 | Sequence 74, Appl |
| 12 | 53.6 | 89.3 | 4766 | 4 | US-09-475-515-73 | Sequence 73, Appl |
| 13 | 43 | 71.7 | 4307 | 4 | US-09-552-950-2 | Sequence 2, Appli |
| 14 | 43 | 71.7 | 9772 | 4 | US-09-552-950-5 | Sequence 5, Appli |
| 15 | 38.2 | 63.7 | 1485 | 4 | US-09-184-418C-73 | Sequence 73, Appl |
| 16 | 38.2 | 63.7 | 9010 | 4 | US-09-184-418C-38 | Sequence 38, Appli |
| 17 | 36 | 60.0 | 1486 | 4 | US-09-184-418C-8 | Sequence 8, Appl |
| 18 | 36 | 60.0 | 8992 | 4 | US-09-184-418C-4 | Sequence 4, Appli |
| 19 | 35 | 58.3 | 1476 | 4 | US-09-184-418C-100 | Sequence 100, App |
| 20 | 35 | 58.3 | 1496 | 4 | US-09-184-418C-82 | Sequence 82, Appl |
| 21 | 35 | 58.3 | 1539 | 4 | US-09-319-588C-3 | Sequence 3, Appli |
| 22 | 35 | 58.3 | 8959 | 4 | US-09-184-418C-11 | Sequence 11, Appl |
| 23 | 35 | 58.3 | 8972 | 4 | US-09-184-418C-9 | Sequence 9, Appli |
| 24 | 35 | 58.3 | 9183 | 4 | US-09-319-588C-1 | Sequence 1, Appli |
| 25 | 35 | 58.3 | 9207 | 3 | US-08-388-353-800 | Sequence 800, App |
| 26 | 35 | 58.3 | 9207 | 3 | US-08-488-551B-800 | Sequence 800, App |
| 27 | 33.4 | 55.7 | 718 | 6 | 5204259-2 | Patent No. 5204259 |
| 28 | 33.4 | 55.7 | 793 | 6 | 5204259-4 | Patent No. 5204259 |
| 29 | 33.4 | 55.7 | 796 | 6 | 5204259-6 | Patent No. 5204259 |
| 30 | 33.4 | 55.7 | 845 | 1 | US-08-589-446-3 | Sequence 3, Appli |
| 31 | 33.4 | 55.7 | 845 | 1 | US-08-444-882-3 | Sequence 3, Appli |
| 32 | 33.4 | 55.7 | 845 | 1 | US-08-389-459A-3 | Sequence 3, Appli |
| 33 | 33.4 | 55.7 | 845 | 3 | US-08-987-867A-3 | Sequence 3, Appli |
| 34 | 33.4 | 55.7 | 871 | 6 | 5204259-10 | Patent No. 5204259 |
| 35 | 33.4 | 55.7 | 1021 | 6 | 5204259-8 | Patent No. 5204259 |
| 36 | 33.4 | 55.7 | 1494 | 4 | US-09-184-418C-47 | Sequence 47, Appl |
| 37 | 33.4 | 55.7 | 1494 | 4 | US-09-184-418C-48 | Sequence 48, Appl |
| 38 | 33.4 | 55.7 | 4307 | 4 | US-09-552-950-1 | Sequence 1, Appli |
| 39 | 33.4 | 55.7 | 4338 | 4 | US-09-872-733A-1 | Sequence 1, Appli |
| 40 | 33.4 | 55.7 | 5362 | 3 | US-09-463-210-5 | Sequence 5, Appli |
| 41 | 33.4 | 55.7 | 5362 | 4 | US-08-850-049-128 | Sequence 128, App |
| 42 | 33.4 | 55.7 | 7228 | 2 | US-08-850-049-128 | Sequence 128, App |
| 43 | 33.4 | 55.7 | 7228 | 2 | US-08-850-049-129 | Sequence 129, App |
| 44 | 33.4 | 55.7 | 7228 | 2 | US-08-050-478-128 | Sequence 128, App |
| 45 | 33.4 | 55.7 | 7228 | 2 | US-08-050-478-129 | Sequence 129, App |

ALIGNMENTS

RESULT 1
US-09-475-515-20
; Sequence 20, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GEBER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 20
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic p55
; OTHER INFORMATION: Gag Major Homology Region
US-09-475-515-20

Query Match      89.3%;  Score 53.6;  DB 4;  Length 60;
Best Local Similarity   93.3%;
Matches  56;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

Qy          1 GACATCAAGCAGGGCCCCAAGGAGGAGGCCCTTCCGCGACTTAGGTGGACCGCGCTTCTTCAAGACC 60
              ||||| || |||||||||| |||||||| |||||||| |||| | ||||||| |||||||| ||||||
Db          1 GACATCCGCCAGGGCCCCCCAAGGAGGAGCCCCTTCCGGCACTACGTGGACCGCTTCTTCACAAGACC 60

RESULT 2
US-09-475-515-9
; Sequence 9, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong

APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 9
LENGTH: 1268
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag

US-09-475-515-9

Query Match             89.3%;  Score 53.6;  DB 4;  Length 1268;
Best Local Similarity   93.3%;  Pred. No. 7.8e-09;
Matches    56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        1 GACATCAAGCACGGCGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 60
            ||||| ||||||| ||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db      862 GACATCCCGCCAGGGGCCCCAAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCTACAAGACC 921

RESULT 3
US-09-475-515-4
Sequence 4, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 4
LENGTH: 1515
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-09-475-515-4

Query Match             89.3%;  Score 53.6;  DB 4;  Length 1515;
Best Local Similarity   93.3%;  Pred. No. 7.9e-09;
Matches    56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        1 GACATCAAGCACGGCGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 60
            ||||| ||||||| ||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db      862 GACATCCCGCCAGGGGCCCCAAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCTACAAGACC 921

RESULT 4
US-09-475-515-5
Sequence 5, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan

---

APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5
LENGTH: 1853
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: HIV-Gag-protease

US-09-475-515-5

Query Match             89.3%;  Score 53.6;  DB 4;  Length 1853;
Best Local Similarity   93.3%;  Pred. No. 8.1e-09;
Matches    56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        1 GACATCAAGCACGGGGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 60
            ||||| |||||| ||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db      862 GACATCCCGCCAGGGGCCCCAAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCTACAAGACC 921

RESULT 5
US-09-475-515-78
Sequence 78, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 78
LENGTH: 1865
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: GP1

US-09-475-515-78

Query Match             89.3%;  Score 53.6;  DB 4;  Length 1865;
Best Local Similarity   93.3%;  Pred. No. 8.1e-09;
Matches    56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        1 GACATCAAGCACGGGGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 60
            ||||| |||||| ||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db      868 GACATCCGCGCCAGGGGCCCCAAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCTACAAGACC 927

RESULT 6
US-09-475-515-79

; Sequence 79, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GP2

Qy

Query Match     89.3%;  Score 53.6;  DB 4;  Length 1865;
Best Local Similarity   93.3%;  Pred. No. 8.1e-09;
Matches   56;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

US-09-475-515-79

Qy        1 GACTCAAGCAGCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTNGGACCCGCTTCTTCTTCAAGACC  60
            |||||  ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      868 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCCGCTTCTTCTTCAAGACC  927

RESULT 7
US-09-475-515-7
Sequence 7, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 2031
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: HIV-Gag/HCV-core fusion polypeptide

Qy

Query Match     89.3%;  Score 53.6;  DB 4;  Length 2031;
Best Local Similarity   93.3%;  Pred. No. 8.1e-09;
Matches   56;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

US-09-475-515-7

Qy        1 GACATCAAGCAGCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCTTCAAGACC  60
            ||||| ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db      862 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGCTTYCTTCTTCAAGACC  921

---

RESULT 8
US-09-475-515-6
Sequence 6, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 4319
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: HIV-Gag-polymerase

Qy

Query Match     89.3%;  Score 53.6;  DB 4;  Length 4319;
Best Local Similarity   93.3%;  Pred. No. 8.8e-09;
Matches   56;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

US-09-475-515-6

Qy        1 GACATCAAGCAGCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCCGCTTCTTCTTCAAGACC  60
            ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      862 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCCGCTTCTTCTTCAAGACC  921

RESULT 9
US-09-475-515-75
Sequence 75, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 75
LENGTH: 4472
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: gp160.modUS4.delV1/V2.gag.modSF2

US-09-475-515-75

Query Match     89.3%;  Score 53.6;  DB 4;  Length 4472;
Best Local Similarity   93.3%;  Pred. No. 8.8e-09;

RESULT 10
US-09-475-515-76
; Sequence 76, Application US/0947551SA
; Patent No. 6602705
; GENERAL INFORMATION:
;   APPLICANT: BARNETT, Susan
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: SRIVASTAVA, Indresh
;   APPLICANT: LIAN, Ying
;   APPLICANT: HARTOG, Karin
;   APPLICANT: LIU, Hong
;   APPLICANT: GREER, Catherine
;   APPLICANT: SELBY, Mark
;   APPLICANT: WALKER, Christopher
;   TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;   TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;   FILE REFERENCE: 1621.002
;   CURRENT APPLICATION NUMBER: US/09/475,515A
;   CURRENT FILING DATE: 1999-12-30
;   NUMBER OF SEQ ID NOS: 90
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 76
;   LENGTH: 4608
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: gp160.modSF162.delV2.gag.modSF2
US-09-475-515-76

Query Match             89.3%;  Score 53.6;  DB 4;  Length 4608;
Best Local Similarity   93.3%;  Pred. No. 8.8e-09;
Matches    56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         1 GACATCAAGCAGGAGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3794 GACATCCGCCAGGGCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 3853

Matches    56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

RESULT 11
US-09-475-515-74
; Sequence 74, Application US/0947551SA
; Patent No. 6602705
; GENERAL INFORMATION:
;   APPLICANT: BARNETT, Susan
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: SRIVASTAVA, Indresh
;   APPLICANT: LIAN, Ying
;   APPLICANT: HARTOG, Karin
;   APPLICANT: LIU, Hong
;   APPLICANT: GREER, Catherine
;   APPLICANT: SELBY, Mark
;   APPLICANT: WALKER, Christopher
;   TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;   TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;   FILE REFERENCE: 1621.002
;   CURRENT APPLICATION NUMBER: US/09/475,515A
;   CURRENT FILING DATE: 1999-12-30
;   NUMBER OF SEQ ID NOS: 90
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 74
;   LENGTH: 4689
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:

OTHER INFORMATION: gp160.modSF162.gag.modSF2
US-09-475-515-74

Query Match             89.3%;  Score 53.6;  DB 4;  Length 4689;
Best Local Similarity   93.3%;  Pred. No. 8.9e-09;
Matches    56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         1 GACATCAAGCAGGAGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3930 GACATCCGCCAGGGCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 3989

RESULT 12
US-09-475-515-73
; Sequence 73, Application US/0947551SA
; Patent No. 6602705
; GENERAL INFORMATION:
;   APPLICANT: BARNETT, Susan
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: SRIVASTAVA, Indresh
;   APPLICANT: LIAN, Ying
;   APPLICANT: HARTOG, Karin
;   APPLICANT: LIU, Hong
;   APPLICANT: GREER, Catherine
;   APPLICANT: SELBY, Mark
;   APPLICANT: WALKER, Christopher
;   TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;   TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;   FILE REFERENCE: 1621.002
;   CURRENT APPLICATION NUMBER: US/09/475,515A
;   CURRENT FILING DATE: 1999-12-30
;   NUMBER OF SEQ ID NOS: 90
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 73
;   LENGTH: 4766
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: gp160.modUS4.gag.modSF2
US-09-475-515-73

Query Match             89.3%;  Score 53.6;  DB 4;  Length 4766;
Best Local Similarity   93.3%;  Pred. No. 8.9e-09;
Matches    56;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         1 GACATCAAGCAGGAGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4011 GACATCCGCCAGGGCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 4070

RESULT 13
US-09-552-950-2
; Sequence 2, Application US/09552950
; Patent No. 6541248
; GENERAL INFORMATION:
;   APPLICANT: Oxford Biomedica (UK) Limited
;   TITLE OF INVENTION: Anti-Viral Vectors
;   FILE REFERENCE: 674524-2004
;   CURRENT APPLICATION NUMBER: US/09/552,950
;   CURRENT FILING DATE: 2000-04-20
;   NUMBER OF SEQ ID NOS: 22
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 2
;   LENGTH: 4307
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:gagpol-SYNgp - codon
;   OTHER INFORMATION: optimised gagpol sequence
US-09-552-950-2

Query Match             71.7%;  Score 43;  DB 4;  Length 4307;

Qy         1 GACATCAAGCAGGAGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 60
             ||||||                   ||||||||||||||||||||||||||||||| ||||||
Db      4088 GACATCCGCCAGGGCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACC 4147

Best Local Similarity 83.1%; Pred. No. 2.5e-05;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GACATCAAGCAGGAGGGGCCCAAGGAGCCCTTCCGCGACTAGCGTGGGACCGCCTTCTTCAAGAC 59
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 850 GACATCCGCCAAGGCCCGAAGGAACCCTTTCGCGACTACGTGGACCGGTTCTTACAAAAC 908

RESULT 14
US-09-552-950-5
; Sequence 5, Application US/09552950
; Patent No. 6541248
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; TITLE OF INVENTION: Anti-Viral Vectors
; FILE REFERENCE: 674524-2004
; CURRENT APPLICATION NUMBER: US/09/552,950
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 5
; LENGTH: 9772
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:pSYNGP
US-09-552-950-5

Qy 1 GACATCAAGCAGGAGGGGCCCAAGGAGCCCTTCCGCGACTAGCGTGGACCGCCTTCTTCAAGAC 59
     ||| ||| |||| ||||||||||| ||||||||||||| |||||||||||| ||| ||||| ||
Db 1957 GACATCCGCCAAGGCCCGAAGGAACCCTTTCGCGACTACGTGGACCGGTTCTTACAAAAC 2015

Query Match 71.7%; Score 43; DB 4; Length 9772;
Best Local Similarity 83.1%; Pred. No. 2.7e-05;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

RESULT 15
US-09-184-418C-73
; Sequence 73, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
;
; SEQ ID NO 73
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=96ZM651.8; gene=gag
US-09-184-418C-73

Qy 1 GACATCAAGCAGGAGGGGCCCAAGGAGCCCTTCCGCGACTACGTGGGACCGGCTTCTTCAAGAC 59
     |||||| ||  || ||||||||| || ||||||||| |||||||||||||| ||||| ||||| ||
Db 847 GACATAAACAAGGCCCAAAGGGAACCCTTTAGAGACTATGTGAGACCGGTTCTTCAAAAC 905

Query Match 63.7%; Score 38.2; DB 4; Length 1485;
Best Local Similarity 78.0%; Pred. No. 0.00081;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Search completed: May 28, 2004, 13:26:22
Job time : 7.49156 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 28, 2004, 08:01:50 ; Search time 234.634 Seconds
(without alignments)
7636.275 Million cell updates/sec

Title: US-09-475-704A-1
Perfect score: 60
Sequence: 1 gactcaagcagggccccaa..........tggaccgcttcttcaagacc 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27512289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
listing first 45 summaries

Databases :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estom:*
16: em_estun:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

SUMMARIES

| Result No. | Score | Query Match Length | DB | ID | Description |
|---|---|---|---|---|---|
| 1 | 29.8 | 49.7 | 686 | 14 | CA180823 | CA180823 SCACST3316 |
| 2 | 29.6 | 49.3 | 568 | 6 | AL809870 | AL809870 Triticum |
| 3 | 29.6 | 49.3 | 1151 | 13 | BQ619287 | BQ619287 RNOSEQ5H1 |
| 4 | 29.6 | 49.3 | 1151 | 13 | BQ619288 | BQ619288 RNOSEQ6A0 |

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| | | | | |
|---|---|---|---|---|
| 5 | 29.6 | 49.3 | 1151 | 13 | BQ619295 RNOSEQ6A0 |
| 6 | 29.6 | 49.3 | 1151 | 13 | BQ619376 RNOSEQ7A0 |
| 7 | 29.6 | 49.3 | 1151 | 13 | BQ619455 RNOSEQ7H0 |
| 8 | 29.6 | 49.3 | 1151 | 13 | BQ619551 CIIIastPC |
| 9 | 29.6 | 49.3 | 1161 | 14 | CK166030 FGAS05007 |
| 10 | 29.2 | 48.7 | 708 | 13 | BU645917 11120340 |
| 11 | 29.2 | 48.7 | 922 | 14 | CK165296 FGAS04924 |
| 12 | 29 | 48.3 | 781 | 14 | CF430199 PH1_26_D0 |
| 13 | 28.6 | 47.7 | 455 | 6 | AL811116 Triticum |
| 14 | 28.6 | 47.7 | 491 | 9 | AU031268 AU031268 |
| 15 | 28.6 | 47.7 | 513 | 14 | CA707887 wdk2c_pk0 |
| 16 | 28.6 | 47.7 | 582 | 6 | AL811117 Triticum |
| 17 | 28.6 | 47.7 | 875 | 14 | CK164904 FGAS04883 |
| 18 | 28.4 | 47.3 | 560 | 12 | BG517417 947062C12 |
| 19 | 28.2 | 47.0 | 408 | 14 | CB639810 OSjNEa12B |
| 20 | 28.2 | 47.0 | 576 | 29 | CB639810 OSjNEa12B |
| 21 | 28.2 | 47.0 | 576 | 29 | CG225792 OGoCF27TH |
| 22 | 28.2 | 47.0 | 585 | 10 | CG225799 OGoCF27TV |
| 23 | 28.2 | 47.0 | 597 | 29 | AW147097 70701LF07 |
| 24 | 28.2 | 47.0 | 625 | 14 | CD884980 F1.118H11 |
| 25 | 28.2 | 47.0 | 646 | 29 | CC675574 OGoAX28TV |
| 26 | 28.2 | 47.0 | 675 | 29 | CC821047 FBCP4B6_U |
| 27 | 28.2 | 47.0 | 809 | 14 | CG249691 OGXCO69TH |
| 28 | 28.2 | 47.0 | 819 | 29 | CK198615 FGAS00710 |
| 29 | 28.2 | 47.0 | 819 | 29 | CG293655 OG0BV73TH |
| 30 | 28.2 | 47.0 | 823 | 29 | CG349088 OGYBL67TV |
| 31 | 28.2 | 47.0 | 829 | 29 | CG288932 OGYCL45TV |
| 32 | 28.2 | 47.0 | 829 | 29 | CC682396 OGLBB77TV |
| 33 | 28.2 | 47.0 | 829 | 29 | CG313210 OGXDR22TV |
| 34 | 28.2 | 47.0 | 858 | 29 | CG364851 OGiAE85TV |
| 35 | 28.2 | 47.0 | 869 | 29 | CC707427 OGUBB52TV |
| 36 | 28.2 | 47.0 | 923 | 29 | CC731054 OGVBI45TV |
| 37 | 28 | 46.7 | 120 | 14 | CD955986 SBX_233_G |
| 38 | 28 | 46.7 | 168 | 10 | BG050418 FM1_53_H0 |
| 39 | 28 | 46.7 | 216 | 9 | AA143927 zEST00762 |
| 40 | 28 | 46.7 | 219 | 10 | BB356040 DG1_121_H |
| 41 | 28 | 46.7 | 262 | 14 | CF061008 QCT16d08_ |
| 42 | 28 | 46.7 | 282 | 12 | BI992820 1020068G0 |
| 43 | 28 | 46.7 | 305 | 13 | BQ657317 HA08N16u |
| 44 | 28 | 46.7 | 306 | 12 | BJ240990 BJ240990 |
| 45 | 28 | 46.7 | 327 | 10 | BF317794 OVl_9_E10 |
| | 28 | 46.7 | 328 | 14 | CA999329 S234G_C07 |

ALIGNMENTS

RESULT 1
CA180823

LOCUS       CA180823                686 bp    mRNA    linear
DEFINITION  SCACST3160C06.g ST3 Saccharum officinarum cDNA clone SCACST3160C06
            5', mRNA sequence.
ACCESSION   CA180823
VERSION     CA180823.1  GI:35114987
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE   1  (bases 1 to 686)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
  COMMENT   Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bccenter.fcav.unesp.br

FEATURES
source

```
          Plate: 160 row: C column: 06
          Seq primer: T7 Promoter Primer.
```

Location/Qualifiers
1..686
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCACST3160C06"
/lab_host="DH10B"
/clone_lib="ST3"
/note="Organ: Fourth apical stalk internodes of adult plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [fourth apical stalk internodes of adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

```
Query Match    49.7%;  Score 29.8;  DB 14;  Length 686;
Best Local Similarity   70.2%;  Pred. No. 5.5e+02;
Matches  40;  Conservative  0;  Mismatches  17;  Indels  0;  Gaps  0;

Qy   1 GACACTCAAGCACAGGGGCCCCCAAGGAGAGCCCTTCCCGGACTACGTGGACCGGCTTCTTCAAGACC 57
        ||| |||||||||||  |||||||||||||| | ||
Db  11 GCCATGACACCCGGACCTCAAGGCCCGGCGGTCACCTCCGCCGACCTGAGAGCTTCCTCAAG 67
```

---

**RESULT 2**

```
ID   AL809870    standard; mRNA; EST; 568 BP.
AC   AL809870;
SV   AL809870.1
DT   05-JUL-2002 (Rel. 72, Created)
DT   02-SEP-2002 (Rel. 72, Last updated, Version 2)
DE   Triticum aestivum (mercia) clone B06_d15_plate_2 embryo (micropylar tip) 6
DE   days post anthesis
KW   EST; expressed sequence tag.
OS   Triticum aestivum (bread wheat)
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
RN   [1]
RA   Wilson I., Beswick R., Shepherd S., Barker G., Parker J., Owen P.,
RA   Edwards D., Coghill J., Holdsworth M., Lenton J., Shewry P., Edwards K.;
RT   "A BBSRC-funded wheat EST resource for the academic community";
RL   Unpublished.
RN   [2]
RP   1-568
RA   Barker G.;
RT   ;
RL   Submitted (03-JUL-2002) to the EMBL/GenBank/DDBJ databases.
RL   Barker G., Institute of Arable Crop Research, Long Ashton, Bristol BS41 9AF
RL   United Kingdom.
FH   Key             Location/Qualifiers
FT   source          1..568
FT                   /db_xref="taxon:4565"
FT                   /mol_type="mRNA"
FT                   /organism="Triticum aestivum"
FT                   /cultivar="mercia"
FT                   /dev_stage="6 days post anthesis"
FT                   /tissue_type="embryo (micropylar tip)"
FT                   /clone_lib="d:15"
FT                   /clone="B06_d15_plate_2"
XX
SQ   Sequence 568 BP; 124 A; 158 C; 194 G; 91 T; 1 other;
```

```
Query Match    49.3%;  Score 29.6;  DB 6;  Length 568;
Best Local Similarity   68.3%;  Pred. No. 6e+02;
Matches  41;  Conservative  0;  Mismatches  19;  Indels  0;  Gaps  0;

Qy   1 GACACTCAAGCACAGGGGCCCCCAAGGAGAGCCCTTCCCGGACTACGTGGACCGGCTTCTTCAAGACC 60
         | || |||| ||||| ||||||| || |
Db 484 GTCATCAAGGACGACCGCCGCCGAGACCTTCCACGACCTCGACGACTACCTCCTCAAGGAAC 543
```

---

**RESULT 3**

```
LOCUS       BQ619287                1151 bp    mRNA    linear   EST 06-SEP-2002
DEFINITION  RNOSEQ5H12_SK.ab1 Salt stressed zea mays roots cDNA library Zea
            mays cDNA clone RNOSEQ5H12_SK.ab1 similar to (AF287276) chlorophyll
            a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION   BQ619287
VERSION     BQ619287.1  GI:21621281
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1151)
  AUTHORS   Wang,H. and Bohnert,H.J.
  TITLE     Genomics of plant stress tolerance
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Mark Fredricksen
            Department of Plant Biology
            University of Illinois
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 2172655473
            Email: bohnertlab@life.uiuc.edu.
FEATURES             Location/Qualifiers
     source          1..1151
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="RNOSEQ5H12_SK.ab1"
                     /tissue_type="Roots"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed zea mays roots cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150
                     mM NaCl"
ORIGIN
```

```
Query Match    49.3%;  Score 29.6;  DB 13;  Length 1151;
Best Local Similarity   73.1%;  Pred. No. 7.1e+02;
Matches  38;  Conservative  0;  Mismatches  14;  Indels  0;  Gaps  0;

Qy   9 GCAGGGCCCCCAAGGAGAGCCCTTCCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60
         ||||  |  | ||| | | |
Db 524 GCAGGACCATCAAGGACCCCGTCAGCGTCACCAGAGACCCCCATCTTCAAGAGAGC 575
```

---

**RESULT 4**

```
LOCUS       BQ619288                1151 bp    mRNA    linear   EST 06-SEP-2002
DEFINITION  RNOSEQ6A01_SK.ab1 Salt stressed zea mays roots cDNA library Zea
            mays cDNA clone RNOSEQ6A01_SK.ab1 similar to (AF287276) chlorophyll
            a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION   BQ619288
VERSION     BQ619288.1  GI:21621282
KEYWORDS    EST.
```

SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1151)
  AUTHORS   Wang,H. and Bohnert,H.J.
  TITLE     Genomics of plant stress tolerance
  JOURNAL   Unpublished (2002)
  COMMENT   Contact: Mark Fredricksen
            Department of Plant Biology
            University of Illinois
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 2172655473
            Email: bohnertlab@life.uiuc.edu.
FEATURES             Location/Qualifiers
     source          1..1151
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="RNOSBQ6A01_SK.ab1"
                     /tissue_type="Roots"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed Zea mays roots cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150
                     mM NaCl"
ORIGIN

Qy          9  GCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC  60
               |||||  |  ||||||||||||| | ||||| ||| ||||||||||||||||||||||
Db        524  GCAGGACATCAAGAACCCCGTCAGGGTCAACCAGGACCCCATCTTCAAGAGC  575

RESULT 5
BQ619295
LOCUS       BQ619295                    1151 bp    mRNA    linear   EST 06-SEP-2002
DEFINITION  RNOSBQ6A08_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
            mays cDNA clone RNOSBQ6A08_SK.ab1 similar to (AF287276) chlorophyll
            a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION   BQ619295
VERSION     BQ619295.1  GI:21621289
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1151)
  AUTHORS   Wang,H. and Bohnert,H.J.
  TITLE     Genomics of plant stress tolerance
  JOURNAL   Unpublished (2002)
  COMMENT   Contact: Mark Fredricksen
            Department of Plant Biology
            University of Illinois
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 2172655473
            Email: bohnertlab@life.uiuc.edu.
FEATURES             Location/Qualifiers
     source          1..1151
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="RNOSBQ6A08_SK.ab1"
                     /tissue_type="Roots"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed Zea mays roots cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150

Query Match             49.3%;  Score 29.6;  DB 13;  Length 1151;
Best Local Similarity   73.1%;            Pred. No. 7.1e+02;
Matches    38;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          9  GCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC  60
               |||||  |  ||||||||||||| | ||||| ||| ||||||||||||||||||||||
Db        524  GCAGGACATCAAGAACCCCGTCAGGGTCAACCAGGACCCCATCTTCAAGAGC  575

mM NaCl"

ORIGIN

RESULT 6
BQ619376
LOCUS       BQ619376                    1151 bp    mRNA    linear   EST 06-SEP-2002
DEFINITION  RNOSBQ7A07_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
            mays cDNA clone RNOSBQ7A07_SK.ab1 similar to (AF287276) chlorophyll
            a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION   BQ619376
VERSION     BQ619376.1  GI:21621370
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1151)
  AUTHORS   Wang,H. and Bohnert,H.J.
  TITLE     Genomics of plant stress tolerance
  JOURNAL   Unpublished (2002)
  COMMENT   Contact: Mark Fredricksen
            Department of Plant Biology
            University of Illinois
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 2172655473
            Email: bohnertlab@life.uiuc.edu.
FEATURES             Location/Qualifiers
     source          1..1151
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="RNOSBQ7A07_SK.ab1"
                     /tissue_type="Roots"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed Zea mays roots cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150
                     mM NaCl"

Query Match             49.3%;  Score 29.6;  DB 13;  Length 1151;
Best Local Similarity   73.1%;            Pred. No. 7.1e+02;
Matches    38;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          9  GCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC  60
               |||||  |  ||||||||||||| | ||||| ||| ||||||||||||||||||||||
Db        524  GCAGGACATCAAGAACCCCGTCAGGGTCAACCAGGACCCCATCTTCAAGAGC  575

ORIGIN

RESULT 7
BQ619455
LOCUS       BQ619455                    1151 bp    mRNA    linear   EST 06-SEP-2002
DEFINITION  RNOSBQ7H05_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
            mays cDNA clone RNOSBQ7H05_SK.ab1 similar to (AF287276) chlorophyll
            a/b-binding protein precursor [Hordeum vulgare], mRNA sequence.
ACCESSION   BQ619455
VERSION     BQ619455.1  GI:21621449
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.

REFERENCE   1 (bases 1 to 1151)
AUTHORS     Wang,H. and Bohnert,H.J.
TITLE       Genomics of plant stress tolerance
JOURNAL     Unpublished (2002)
COMMENT     Contact: Mark Fredricksen
            Department of Plant Biology
            University of Illinois
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 2172655473
            Email: bohnertlab@life.uiuc.edu.
FEATURES             Location/Qualifiers
     source          1..1151
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="RNOSEQ7H05_SK.ab1"
                     /tissue_type="Roots"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed zea mays roots cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150
                     mM NaCl"
ORIGIN

Query Match       49.3%;  Score 29.6;  DB 13;  Length 1151;
Best Local Similarity   73.1%;  Pred. No. 7.1e+02;
Matches  38;  Conservative   0;  Mismatches  14;  Indels   0;  Gaps   0;

Qy          9 GCAGGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60
               ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 GCAGGACATCAAGGAACCCCGTCAACCAGGAGACCCCATCTTCTTCAAGAGC 575

RESULT 8
LOCUS       BQ619551
            BQ619551  1151 bp    mRNA    linear   EST 06-SEP-2002
DEFINITION  CillastPCRs_SK.ab1 Salt stressed Zea mays leaves cDNA library Zea
            mays cDNA clone CillastPCRs_SK.ab1 similar to (AF287276)
            chlorophyll a/b-binding protein precursor [Hordeum vulgare], mRNA
            sequence.
ACCESSION   BQ619551
VERSION     BQ619551.1  GI:21621545
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 1151)
AUTHORS     Wang,H. and Bohnert,H.J.
TITLE       Genomics of plant stress tolerance
JOURNAL     Unpublished (2002)
COMMENT     Contact: Mark Fredricksen
            Department of Plant Biology
            University of Illinois
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 2172655473
            Email: bohnertlab@life.uiuc.edu.
FEATURES             Location/Qualifiers
     source          1..1151
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="CillastPCRs_SK.ab1"
                     /tissue_type="Leaves"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed zea mays leaves cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150
                     mM NaCl"
ORIGIN

Query Match       49.3%;  Score 29.6;  DB 13;  Length 1151;

RESULT 9
LOCUS       CK166030
            CK166030  1161 bp    mRNA    linear   EST 05-DEC-2003
DEFINITION  FGAS050078 Triticum aestivum FGAS: Talt7 Triticum aestivum cDNA,
            mRNA sequence.
ACCESSION   CK166030
VERSION     CK166030.1  GI:3899675
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 1161)
AUTHORS     Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,T.D., Laroche,A.,
            Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.I. and Sarhan,F.
TITLE       Functional Genomics of Abiotic Stress in Wheat and Canola Crops
JOURNAL     Unpublished (2003)
COMMENT     Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            IC101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas.ests@cs.usask.ca
            This sequence is the direct result of the Base calling software
            Phred (default parameters). It is the raw base calls. To aid in the
            identification of the high quality insert the software Lucy
            (default parameters) has been run on this sequence. Lucy identified
            the region [62,774].
            Plate: Talt707 row: H column: 16.
FEATURES             Location/Qualifiers
     source          1..1161
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Wheat line CI 14106"
                     /db_xref="taxon:4565"
                     /clone="Organ: Crown; Vector: pGERM-T; SSH (suppression
                     /lab_host="DH5 alpha"
                     /clone_lib="Triticum aestivum FGAS: Talt7"
                     /note="Organ: Crown; Vector: pGERM-T; SSH (suppression
                     subtractive hybridization) cDNA library from genotype
                     CI14106 cold hardened at 2 C for 21 days and 49 days
                     (equal amount of cDNA pooled together before subtraction,
                     tester) and subtracted against genotype CI14106
                     non-hardened (20 C) (driver). Nitro-Pyrole anchored
                     oligo-dT priming and non-directional cloning."
ORIGIN

Query Match       49.3%;  Score 29.6;  DB 14;  Length 1161;
Best Local Similarity   68.3%;  Pred. No. 7.1e+02;
Matches  41;  Conservative   0;  Mismatches  19;  Indels   0;  Gaps   0;

Qy          9 GCAGGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60
               ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 GTCATCAAGGAAGGAGACCCGCCAAGACCTTCCACGACCTCGACGACCACCTCCTCAAGAAC 613

RESULT 10
LOCUS       BU645917
            BU645917  708 bp    mRNA    linear   EST 30-SEP-2002
DEFINITION  1120348O7.x1  C. reinhardtii CC-1690 (mt+), CC-1691 (mt-), Gamete
            (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA

sequence.
ACCESSION   BU645917
VERSION     BU645917.1  GI:23358097
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 708)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,B., Hauser,C.,
            Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1112
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Charles Hauser
            Duke University
            DCMB Box 91000
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1..708
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="21gr (CC-1690 wild type mt+) & 6145c (CC-1691
                     wild type mt-)"
                     /clone_lib="pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; Gamete library was constructed by Hui Zhao, Min Lu,
                     Jeffrey McDermott, William J. Snell and John Davies.
                     strain 21gr cells (CC-1690; mating type plus) and strain
                     6145c cells (CC-1691; mating type minus) that had been
                     growing on a light-dark cycle (13:11 L/D) in R-medium
                     (Sager and Granick) were separately transferred into
                     nitrogen-free medium at 8 hours into the light period.
                     PolyA mRNA was purified from each sample every 2 hours for
                     the next 18 hours. The mRNA was pooled and used for cDNA
                     synthesis. The cDNA was directionally cloned into lambda
                     Zap II (Stratagene) in the EcoRI (5') and XhoRI (3')
                     sites. pBluescript II SK- plasmid were excised from the
                     lambda ZAP clones by superinfection with ExAssist
                     (Stratagene) phage. The library was normalized using
                     method 4 described in Bonaldo et al. (1996) Genome
                     Research 6: 791-806."
                     /clone="C. reinhardtii CC-1690 (mt+), CC-1691 (mt-),
                     Gamete (normalized), Lambda Zap II"
                     /db_xref="taxon:3055"
ORIGIN

RESULT 11
Query Match             48.7%;  Score 29.2;  DB 13;  Length 708;
Best Local Similarity   69.0%;
Matches    40;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy        3 CATCAAGCAGAGGGCCCCCAAGGAGGCCCTTCCCGGACTACGTGGACCGCTTCTTCAAGACC 60
            ||||||  ||| ||  |||| ||  ||| ||||||||| || ||  ||||||  |||||||||||
Db      564 CTTCAACCGCGACCCCATCAACCGCTTCATCCTTACTTCCACCGCTACTTCGTCAAGGCC 621

---

LOCUS       CK165296          922 bp    mRNA    linear   EST 05-DEC-2003
DEFINITION  FGAS049245 Triticum aestivum FGAS: TaIt7 Triticum aestivum cDNA,
            mRNA sequence.
ACCESSION   CK165296
VERSION     CK165296.1  GI:38997204
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 922)
  AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Budes,F., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
            Linke,M.G., McCarthy,B.L., Monroy,A., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE     Functional Genomics of Abiotic Stress in Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas_ests@cs.usask.ca
            This sequence is the direct result of the Base calling software
            Phred (default parameters). It is the raw base calls. To aid in the
            identification of the high quality insert the software Lucy
            (default parameters) has been run on this sequence. Lucy identified
            the region [123,581].
            Plate: TaIt705 row: A column: 03.
FEATURES             Location/Qualifiers
     source          1..922
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Wheat line CI 14106"
                     /db_xref="taxon:4565"
                     /clone_lib="DH5 alpha"
                     /lab_host="taxon:4565"
                     /clone="Organ: Crown; Vector: pGEM-T; SSH (suppression
                     subtractive hybridization) cDNA library from genotype
                     CI14106 cold hardened at 2 C for 21 days and 49 days
                     /note="Organ: Triticum aestivum FGAS: TaIt7"
                     (equal amount of cDNA pooled together before subtraction,
                     tester) and subtracted against genotype CI14106
                     non-hardened (20 C) (driver). Nitro-Pyrole anchored
                     oligo-dT priming and non-directional cloning."
ORIGIN

RESULT 12
Query Match             48.7%;  Score 29.2;  DB 14;  Length 922;
Best Local Similarity   66.7%;
Matches    40;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy        1 GACATCAAGCAGAGGGCCCCCAAGGAGGCCCTTCCCGGACTACGTGGACCGCTTCTTCAAGACC 60
            | ||||||| || ||  ||||| || ||  |||  ||||| || ||  | ||  |||  | |||||| |
Db      615 GTCATCAAGGAGGAGAACCCGCGNCNAGGACCTTCCACGACCTCGGACGACTACCTCCTCAAGGAAC 674

---

LOCUS       CF430199          781 bp    mRNA    linear   EST 03-SEP-2003
DEFINITION  PHI_26_D03.g1_A002 Phosphorous-deficient seedlings Sorghum bicolor
            cDNA clone PHI_26_D03_A002 5', mRNA sequence.
ACCESSION   CF430199
VERSION     CF430199.1  GI:34442900
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 781)
  AUTHORS   Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C.,
            Sun,P., Sullivan,R., Harris,K., Eastman,A. and Pratt,L.H.
  TITLE     An EST database from Sorghum: phosphorous-deficient seedlings
  JOURNAL   Unpublished (2003)
COMMENT     Other ESTs: PHI_26_D03.b1_A002
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in

FH   Key                                                    Location/Qualifiers

FEATURES             Location/Qualifiers
source               1..781
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /cultivar="BTx623"
                     /db_xref="taxon:4558"
                     /clone="PH1_26_D03_A002"
                     /lab_host="DH10B-T1 phage-resistant B. coli"
                     /clone_lib="Phosphorous-deficient seedlings"
                     /note="Vector: pMB18S-FL3; Site 1: XhoI; Site 2: XhoI; The
                     library was prepared from polyA+ RNA from BTx623 seedlings
                     grown hydroponically in the absence of added phosphorous.
                     At 14 days of age, roots and shoots were harvested and
                     stored at -80 C until RNA was isolated. Double-stranded
                     cDNA was cloned unidirectionally into different DraIII
                     sites of the pMB18S-FL3 vector (5-prime DraIII site is
                     CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises
                     the cDNA insert."
ORIGIN

Qy   Seq primer: Sugs (CTTCTGCTCTAAAAGCTCG).

Query Match     48.3%;  Score 29;  DB 14;  Length 781;
Best Local Similarity   71.7%;  Pred. No. 9.3e+02;
Matches   38;  Conservative   0;  Mismatches   15;  Indels   0;  Gaps   0;

Qy        8 AGCAGGGCCCCAAGGGAGCCCTTCCGCGACTAGGTGGACCGGCTTCTTCAAGACC 60
            ||||||||||| |||||||||||||||||||||| ||||||||||||||||||||
Db      580 AGCAGGACATCAAGGAGACCCGGCCAGCGTCAACCAGGAGACCCCATCTTCAAGAGACC 632

───────────────────────────────────────────────────────────

RESULT 13
ID   AL811116
AC   AL811116;
XX
SV   AL811116.1
XX
DT   05-JUL-2002 (Rel. 72, Created)
DT   02-SEP-2002 (Rel. 72, Last updated, Version 2)
XX
DE   Triticum aestivum (mercia) clone A10_d37_plate_1 embryo (micropylar tip) 10
DE   days post anthesis
XX
KW   EST; expressed sequence tag.
XX
OS   Triticum aestivum (bread wheat)
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Triticeae; Triticum.
XX
RN   [1]
RA   Wilson I., Beswick R., Shepherd S., Barker G., Parker J., Owen P.,
RA   Edwards D., Coghill J., Holdsworth M., Lenton J., Shewry P., Edwards K.;
RT   "A BBSRC-funded wheat EST resource for the academic community";
RL   Unpublished.
XX
RN   [2]
RP   1-455
RA   Barker G.;
RT   ;
RL   Submitted (03-JUL-2002) to the EMBL/GenBank/DDBJ databases.
RL   Barker G., Institute of Arable Crop Research, Long Ashton, Bristol BS41 9AF
RL   United Kingdom.
XX
FH   Key                                                    Location/Qualifiers

FEATURES             Location/Qualifiers
source               1..455
                     /db_xref="taxon:4565"
                     /mol_type="mRNA"
                     /organism="Triticum aestivum"
                     /cultivar="mercia"
                     /dev_stage="10 days post anthesis"
                     /tissue_type="embryo (micropylar tip)"
                     /clone_lib="d:37"
                     /clone="A10_d37_plate_1"
XX
SQ   Sequence 455 BP; 95 A; 140 C; 149 G; 70 T; 1 other;

Query Match     47.7%;  Score 28.6;  DB 6;  Length 455;
Best Local Similarity   66.7%;  Pred. No. 1e+03;
Matches   40;  Conservative   0;  Mismatches   20;  Indels   0;  Gaps   0;

Qy        1 GACATCAAGCAGGGCCCCAAGGGAGCCCTTCCGCGACTAGGTGGACCGGCTTCTTCAAGACC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      324 GTCATCAAGGAGAGACCCGGCCCGCCCNAGACCTTCCACGACNTCGACGACTACCTCCTCAAGAAC 383

───────────────────────────────────────────────────────────

RESULT 14
LOCUS       AU031268                 491 bp    mRNA    linear    EST 01-APR-2002
DEFINITION  AU031268 Rice cDNA from immature leaf including apical meristem
            Oryza sativa (japonica cultivar-group) cDNA clone E61244_22, mRNA
            sequence.
ACCESSION   AU031268
VERSION     AU031268.1     GI:3767158
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 491)
AUTHORS     Sasaki,T. and Yamamoto,K.
TITLE       Rice cDNA from immature leaf including apical meristem
JOURNAL     Unpublished (1997)
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT = RGP
            POLYA=No.
FEATURES             Location/Qualifiers
source               1..491
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="E61244_22"
                     /dev_stage="immature"
                     /clone_lib="Rice cDNA from immature leaf including apical
                     meristem"
                     /note="Organ: leaf; immature leaf including apical
                     meristem (under long day condition)"
ORIGIN

Query Match     47.7%;  Score 28.6;  DB 9;  Length 491;
Best Local Similarity   66.7%;  Pred. No. 1.1e+03;
Matches   40;  Conservative   0;  Mismatches   20;  Indels   0;  Gaps   0;

Qy        1 GACATCAAGCAGGGCCCCCAAGGGAGCCCTTCCGCGACTAGGTGGACCGGCTTCTTCAAGACC 60
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      183 GTCATCAAGGAGGACCCGGCCCCAAGACACTTCCACGACNTCGACGACTACCTCCTCAAGAAC 242

───────────────────────────────────────────────────────────

RESULT 15

```
CA707887/c
LOCUS           CA707887                  513 bp    mRNA       linear    EST 26-NOV-2002
DEFINITION      wdk2c.pk007.j12 wdk2c Triticum aestivum cDNA clone wdk2c.pk007.j12
                5' end, mRNA sequence.
ACCESSION       CA707887
VERSION         CA707887.1  GI:25429680
KEYWORDS        EST.
SOURCE          Triticum aestivum (bread wheat)
ORGANISM        Triticum aestivum
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                Pooideae; Triticeae; Triticum.
REFERENCE       1  (bases 1 to 513)
AUTHORS         Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
                Miao,G., Caraher,N. and Hanafey,M.K.
TITLE           DuPont Wheat cDNA Sequence
JOURNAL         Unpublished (2002)
COMMENT         Contact: Scott V. Tingey
                Crop Genetics
                E. I. DuPont de Nemours and Company
                1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
                Tel: 302-631-2602
                Fax: 302-631-2607
                Email: Scott.V.Tingey@USA.dupont.com
                Seq primer: M13.
FEATURES             Location/Qualifiers
     source          1..513
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone="wdk2c.pk007.j12"
                     /tissue_type="kernel"
                     /clone_lib="wdk2c"
                     /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
                     XhoI; Wheat (Triticum aestivum L.) developing kernel, 7
                     days after anthesis."
ORIGIN
Query Match      47.7%;    Score 28.6;  DB 14;   Length 513;
Best Local Similarity  66.7%;   Pred. No. 1.1e+03;
Matches   40;  Conservative   0;  Mismatches  20;  Indels   0;  Gaps  0;

Qy        1  GACATCAAGCAGGGCGCCCCAAGGGAGCGCCTTCCGCGACTTACGTGGACGGCTTCTTCAAGGACC    60
             |||||||| |||||  |||  |||||| ||| || ||||| ||||||| ||||||||||||||
Db      295  GTCATCAAGGAGACCGCCGCGCCAAGACCTTCCACGACTNGAGGACTACCTCCTCAAGAAC    236
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      June  1, 2004, 11:29:09 ; Search time 695 Seconds
             (without alignments)
             9876.679 Million cell updates/sec

Title:       US-09-475-704A-4
Perfect score: 1509
Sequence:    1 atgggcgccgcgcagcat..........gcggccccctgagccagtaa 1509

Scoring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

Searched:    2960401 seqs, 2274450654 residues

Total number of hits satisfying chosen parameters:     5920802

Post-processing: Minimum Match 0%
             Maximum Match 100%
             Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :   Published Applications_NA:*
    1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
    2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
    3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
    4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
    5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
    6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
    7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
    8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
    9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
   10: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
   11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
   12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
   13: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
   14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
   15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
   16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
   17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
   18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
   19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Match Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1509 | 100.0 | 1509 | 13 | US-09-899-575-4 | Sequence 4, Appli |
| 2 | 1507.4 | 99.9 | 1509 | 10 | US-09-967-464-64 | Sequence 64, Appl |
| 3 | 1504.2 | 99.7 | 1509 | 10 | US-10-190-435-17 | Sequence 17, Appl |
| 4 | 1489.8 | 98.7 | 1509 | 13 | US-09-899-575-21 | Sequence 21, Appl |
| 5 | 1297.6 | 86.0 | 1494 | 13 | US-09-899-575-51 | Sequence 51, Appl |
| 6 | 1294.8 | 85.8 | 1491 | 13 | US-09-899-575-99 | Sequence 99, Appl |
| 7 | 1276.8 | 84.6 | 1479 | 10 | US-09-967-464-63 | Sequence 63, Appl |
| 8 | 1276.8 | 84.6 | 1479 | 13 | US-09-899-575-3 | Sequence 3, Appli |
| 9 | 1275.2 | 84.5 | 3162 | 15 | US-09-899-575-2 | Sequence 2, Appli |
| 10 | 1275.2 | 84.5 | 3462 | 15 | US-10-190-435-18 | Sequence 18, Appl |
| 11 | 1275.2 | 84.5 | 4419 | 15 | US-10-190-435-19 | Sequence 19, Appl |
| 12 | 1275.2 | 84.5 | 4419 | 15 | US-10-190-305A-14 | Sequence 14, Appl |
| 13 | 1275.2 | 84.5 | 4483 | 15 | US-10-190-435-35 | Sequence 35, Appl |
| 14 | 1275.2 | 84.5 | 4606 | 15 | US-10-190-435-34 | Sequence 34, Appl |
| 15 | 1275.2 | 84.5 | 4702 | 15 | US-10-190-435-36 | Sequence 36, Appl |
| 16 | 1275.2 | 84.5 | 4702 | 15 | US-10-190-435-38 | Sequence 38, Appl |
| 17 | 1275.2 | 84.5 | 4716 | 15 | US-10-190-435-17 | Sequence 17, Appl |
| 18 | 1275.2 | 84.5 | 4716 | 15 | US-10-190-305A-13 | Sequence 13, Appl |
| 19 | 1273.6 | 84.4 | 2742 | 15 | US-10-190-435-20 | Sequence 20, Appl |
| 20 | 1273.6 | 84.4 | 2742 | 15 | US-10-190-305A-15 | Sequence 15, Appli |
| 21 | 1273.6 | 84.4 | 3930 | 15 | US-10-190-435-9 | Sequence 9, Appl |
| 22 | 1273.6 | 84.4 | 3930 | 15 | US-10-190-305A-9 | Sequence 9, Appl |
| 23 | 1273.6 | 84.4 | 3930 | 15 | US-10-190-435-10 | Sequence 10, Appl |
| 24 | 1273.6 | 84.4 | 5145 | 15 | US-10-190-435-11 | Sequence 11, Appl |
| 25 | 1273.6 | 84.4 | 5145 | 15 | US-10-190-305A-12 | Sequence 12, Appl |
| 26 | 1272.6 | 84.3 | 4713 | 15 | US-10-190-435-59 | Sequence 59, Appl |
| 27 | 1272.6 | 84.3 | 4713 | 15 | US-10-190-305A-83 | Sequence 83, Appl |
| 28 | 1271.6 | 84.3 | 5184 | 15 | US-10-190-305A-58 | Sequence 58, Appl |
| 29 | 1271.6 | 84.3 | 5184 | 15 | US-10-190-305A-82 | Sequence 82, Appl |
| 30 | 1271.2 | 84.2 | 1479 | 10 | US-09-967-464-67 | Sequence 67, Appl |
| 31 | 1270.2 | 84.2 | 2742 | 15 | US-10-190-435-57 | Sequence 57, Appl |
| 32 | 1270.2 | 84.2 | 2742 | 15 | US-09-899-575-81 | Sequence 81, Appl |
| 33 | 1260.8 | 83.6 | 1479 | 15 | US-10-190-305A-81 | Sequence 81, Appl |
| 34 | 1201 | 79.6 | 1515 | 13 | US-10-387-336-4 | Sequence 4, Appl |
| 35 | 1201 | 79.6 | 4472 | 13 | US-10-387-336-75 | Sequence 75, Appl |
| 36 | 1201 | 79.6 | 4608 | 13 | US-10-387-336-76 | Sequence 76, Appl |
| 37 | 1201 | 79.6 | 4689 | 13 | US-10-387-336-74 | Sequence 74, Appl |
| 38 | 1201 | 79.6 | 4766 | 13 | US-10-387-336-73 | Sequence 73, Appl |
| 39 | 1199.4 | 79.5 | 2799 | 15 | US-10-241-009-18 | Sequence 18, Appl |
| 40 | 1199.4 | 79.5 | 2799 | 15 | US-10-387-336-18 | Sequence 18, Appl |
| 41 | 1199.4 | 79.5 | 2799 | 13 | US-10-190-434B-18 | Sequence 18, Appl |
| 42 | 1199.4 | 79.5 | 3205 | 15 | US-10-190-305A-11 | Sequence 11, Appl |
| 43 | 1199.4 | 79.5 | 3205 | 15 | US-10-190-434B-17 | Sequence 17, Appl |
| 44 | 1199.4 | 79.5 | 3496 | 15 | US-10-241-009-15 | Sequence 15, Appl |
| 45 | 1199.4 | 79.5 | 3496 | 15 | US-10-190-434B-15 | Sequence 15, Appl |

ALIGNMENTS

RESULT 1
US-09-899-575-4
; Sequence 4, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 4
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: of HIV strain AF110967
US-09-899-575-4

Query Match          100.0%;  Score 1509;  DB 13;  Length 1509;
Best Local Similarity  100.0%;  Pred. No. 0;  Matches
Matches 1509;  Conservative  0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy         1 ATGGGCGCCGCGCAGCATGTCAGTGGGGGGAAGAGCACTACATGCTGACGCAACCTGGGTTGGACAAGTGGGAGAAGATCCGC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGGGCGCCGCGCAGCATGTCAGTGGGGGGAAGAGCACTACATGCTGACGCAACCTGGGTTGGACAAGTGGGAGAAGATCCGC 60

Qy        61 CTGCGCCCGGGCGGCAAGAAGCACTACATGCTGAAGCACATCGTGTGGGCCCAGCGCCGGGCGAG 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CTGCGCCCGGGCGGCAAGAAGCACTACATGCTGAAGCACATCGTGTGGGCCCAGCGCCGGGCGAG 120

RESULT 2
US-09-967-464-64
; Sequence 64, Application US/09967464
; Publication No. US20030184453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Manmohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: PP16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 64
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-64

Query Match          99.9%;  Score 1507.4;  DB 10;  Length 1509;
Best Local Similarity   99.9%;
Matches 1508;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 ATGGGCGGCCGCGCGCCCAGCATCTCTGCGCGGCGGAGAAGTTGGAGAAGAGATCCGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGGGCGGCCGCGCGCCCAGCATCTCTGCGCGGCGGAGAAGTTGGAGAAGAGATCCGC 60

Qy       61 CTCGCGCCCCCGGCGCAAGAGACTACATGCTGAAGCACTTGGTGTGTGGGGCCAGCGCCGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CTCGCGCCCCCGGCGCAAGAGACTACATGCTGAAGCACTTGGTGTGTGGGGCCAGCGCCGAG 120

Qy      121 CTGGAGGGGCTTCGCCCCTGAACCCCGGAGGCTTCAAGCAGCGATC 180
            ||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CTGGAGGGGCTTCGCCCCTGAACCCCGGAGGCTTCAAGCAGCGATC 180

Qy      181 ATGAAGCAGCTGCAGCCCCCTGAGAC... 240
            |||||||||||||||||||||||||||
Db      181 ATGAAGCAGCTGCAGCCCCCTGAGAC... 240

Qy     1141 GGCCCCCGGCCAACCTCAAGTGCTTCAACTGCGGGAAGGAGGGCCACATCGCCCGAGAAC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 GGCCCCCGGCCAACCTCAAGTGCTTCAACTGCGGGAAGGAGGGCCACATCGCCCGAGAAC 1200

Qy     1501 AGCCAGTAA 1509
            |||||||||
Db     1501 AGCCAGTAA 1509

```
Qy    181 ATGAAGCAGCTGCACCGGAGCACCCGGCACCGGAGCGAGAGCGAGCTGCGCGACAAC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATGAAGCAGCTGCACCGGAGCACCCGGCACCGGAGCGAGAGCGAGCTGCGCGACAAC 240

Qy    241 ACCGGTGGCCACCCTGTACTGGGTGCACGGCGGCCATCGATGAGGTCCGGACACCAAGGAGGCC 300
Db    241 ACCGGTGGCCACCCTGTACTGGGTGCACGGCGGCCATCGATGAGGTCCGGACACCAAGGAGGCC 300
```

```
Qy      1 ATGGGGCGCCCCGCCCCAGCATCCTGCGCGGGCGGAGAGCTGGACAAGTGGGAGAGATCCGC 60
Db      1 ATGGGGCGCCCCGCCCCAGCATCCTGCGCGGGCGGAGAGCTGGACAAGTGGGAGAGATCCGC 60
```

RESULT 4
US-09-899-575-21

; Sequence 21, Application US/0989575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
;  APPLICANT: Zur Megede, Jan
;  APPLICANT: Barnett, Susan W.
;  APPLICANT: Egnelbrecht, Susan
;  APPLICANT: van Rensburg, Estrelita Janse
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP01631.102
;  CURRENT APPLICATION NUMBER: US/09/899,575
;  CURRENT FILING DATE: 2001-07-05
;  PRIOR APPLICATION NUMBER: 09/475,704
;  PRIOR FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 21
;  LENGTH: 1509
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: Synthetic Gag
;  OTHER INFORMATION: coding sequence of HIV strain AF110967
US-09-899-575-21

Query Match        98.7%;  Score 1689.8;  DB 13;  Length 1509;
Best Local Similarity  99.2%;  Pred. No. 1.7e-311;
Matches 1497;  Conservative 0;  Mismatches 12;  Indels 0;  Gaps 0;

```
Qy      1  ATGGGCGCGCCCGGCCCAGCAATCCTGCGCGGCGGAAGCTGGAAGAAGATCCCGC  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGGGCGCGCCCGGCCCAGCATCCTGCGCGGCGGAAGCTGGACAAGTGGGAGAAGATCCGC  60

Qy     61  CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG  120
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG  120

Qy    121  CTGGAGCGGTTCGCCCTGAACCCCGGCCTGCTGGAGACCCCGGAGGGCTGCAAGCAGATC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  CTGGAGCGGTTCGCCCTGAACCCCGGCCTGCTGGAGACCCCGGAGGGCTGCAAGCAGATC  180

Qy    181  ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCTCCTTGTACAAC  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCTCCTTGTACAAC  240

Qy    241  ACCGTGGCCACCCTGTACTGCGTGCACGCCGAGATCCCCGACACCAAGGAGGCC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  ACCGTGGCCACCCTGTACTGCGTGCACGCCGAGATCCCCGACACCAAGGAGGCC  300
```

(Alignment continues Qy/Db rows from positions 301 through 1509 / 1500.)

```
Qy   1441  CGCGGAAGCGCCTACCGCGAGCCCCTGACCCGCCCTGTTCGGCAGCAGCGGCCCCTG  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  CGCGGAGCCCTACCTGCGAGCCCCTGACCCGCCCTGTTCGGCAGCAGCGGCCCCTG  1500

Qy   1501  AGCCAGTAA  1509
           |||||||||
Db   1501  AGCCAGTAA  1509
```

RESULT 5
US-09-899-575-51
; Sequence 51, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
;  APPLICANT: Zur Megede, Jan
;  APPLICANT: Barnett, Susan W.
;  APPLICANT: Egnelbrecht, Susan
;  APPLICANT: van Rensburg, Estrelita Janse
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP01631.102
;  CURRENT APPLICATION NUMBER: US/09/899,575
;  CURRENT FILING DATE: 2001-07-05
;  PRIOR APPLICATION NUMBER: 09/475,704
;  PRIOR FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 51
;  LENGTH: 1494
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: HIV Type C
;  OTHER INFORMATION: Gag optimized
US-09-899-575-51

Query Match         86.0%;  Score 1297.6;  DB 13;  Length 1494;
Best Local Similarity   92.7%;  Pred. No. 3.9e-270;
Matches 1402;  Conservative   0;  Mismatches  89;  Indels  21;  Gaps  3;

RESULT 6
US-09-899-575-99
; Sequence 99, Application US/0899575
; Publication No. US20030229961A1
; GENERAL INFORMATION:
APPLICANT: Zur Megede, Jan
APPLICANT: Barnett, Susan W.
APPLICANT: Bgnelbrecht, Susan
APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 99
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gag_TV2_C_ZAopt
US-09-899-575-99

Query Match        85.8%;   Score 1294.8;  DB 13;  Length 1491;
Best Local Similarity   92.7%;   Pred. No. 1.6e-269;
Matches 1401;  Conservative   0;  Mismatches   87;  Indels  24;  Gaps   3;

RESULT 7
US-09-967-464-63
Sequence 63, Application US/09967464
Publication No. US20030138453A1
GENERAL INFORMATION:
 APPLICANT: O'Hagan, Derek
 APPLICANT: Otten, Gillis
 APPLICANT: Donnelly, John J.
 APPLICANT: Polo, John M.
 APPLICANT: Barnett, Susan
 APPLICANT: Singh, Manmohan
 APPLICANT: Ulmer, Jeffrey
 APPLICANT: Dubensky, Jr., Thomas W.
 TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
 FILE REFERENCE: PP16269.004
 CURRENT APPLICATION NUMBER: US/09/967,464
 CURRENT FILING DATE: 2002-04-11
 PRIOR APPLICATION NUMBER: 60/236,105
 PRIOR FILING DATE: 2000-09-28
 PRIOR APPLICATION NUMBER: 2000-09-28
 PRIOR APPLICATION NUMBER: 60/315,905
 PRIOR FILING DATE: 2001-08-30

; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-63

Query Match    84.6%;  Score 1276.8;  DB 10;  Length 1479;
Best Local Similarity   92.5%;  Pred. No. 1.2e-265;
Matches 1399;  Conservative   0;  Mismatches 77;  Indels 36;  Gaps 4;

[Sequence alignment — Qy (Query) and Db (Database) paired rows with midline match bars, positions 1–1509 (Db) / 1–957 (Qy), continuing across the page]

```
              ||||||||||||||||||  ||||||||| |||||||||||||||||||| |||||||||
Db        901 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960

Qy        958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCACC 1017
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGC 1020

Qy       1018 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1077
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080

Qy       1078 CTGGCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAGAAGAGCAACTTC 1137
              |||||||||||||||||||||||||||| |   | | | |||||||||||||||||||||
Db       1081 CTGGCCGAGGCGATGAGCCAGGCCAACACC---AGCGTGATGATGCAGAAGAGCAACTTC 1137

Qy       1138 AAGGGCCCCCGGCGCAACGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAG 1197
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1138 AAGGGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGC 1197

Qy       1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257

Qy       1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1317

Qy       1318 CGCCCCGGCAACTTCCTGCAGAACCGCAGCGAGCCCGCCGCCCCCACCGTGCCCACCGCC 1377
              |||||||||||||||||||||| |||                  | ||| ||||||||||
Db       1318 CGCCCCGGCAACTTCCTGCAGAGCCG------------------CCCCGAGCCCACCGCC 1359

Qy       1378 CCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGCCCCCAAGCAGGAGCCCAAG 1437
              |||||||||||||||||||||||||||||||||||||||| ||  |||||||||  ||||
Db       1360 CCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGGCCAGAAGCAGGAGAGCAAG 1419

Qy       1438 GACCGCGAGCCCTACCGCGAGCCCCTGACCGCCCTGCGCAGCCTGTTCGGCAGCGGCCCC 1497
              |            |||||||| ||||||||  ||||   ||||||||||||| || ||||
Db       1420 G------------ACCGCGAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCC 1467

Qy       1498 CTGAGCCAGTAA 1509
              ||||||||||||
Db       1468 CTGAGCCAGTAA 1479

RESULT 8
US-09-899-575-3
; Sequence 3, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
;  APPLICANT: Zur Megede, Jan
;  APPLICANT:  Barnett, Susan W.
;  APPLICANT:  Egnelbrecht, Susan
;  APPLICANT:  van Rensburg, Estrelita Janse
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION:  POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP01631.102
;  CURRENT APPLICATION NUMBER: US/09/899,575
;  CURRENT FILING DATE:  2001-07-05
;  PRIOR APPLICATION NUMBER: 09/475,704
;  PRIOR FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1479
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
;   OTHER INFORMATION: of HIV strain AF110965
US-09-899-575-3

  Query Match             84.6%;  Score 1276.8;  DB 13;  Length 1479;
  Best Local Similarity   92.5%;  Pred. No. 1.2e-265;
  Matches 1399;  Conservative    0;  Mismatches   77;  Indels   36;  Gaps    4;

Qy          1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGAGAAGCTGGACAAGTGGGAGAAGATCCGC 60
              |||||||||||||||||||||||||||||||  |||||||||   ||||||   ||||||
Db          1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60

Qy         61 CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
              |||||||||||||||||||||  ||||||| |||||||||||||||||||||||||||||
Db         61 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120

Qy        121 CTGGAGGGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCAAGCAGATC 180
              ||||||   ||||||||||||||||||||||||||||||  |||||||||||||||||||
Db        121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180

Qy        181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC 240
              ||    |||||||| ||||||||||||||||||| ||||||||||   ||||||| ||||
Db        181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240

Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTCCGCGACACCAAGGAGGCC 300
              ||||||||||||||||||||||||||||     |||||||||||||||||||||||||||
Db        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAG 360
              ||||||||||||||||||||||||||||||| ||||||||||| |||||||||| |||
Db        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360

Qy        361 GCCGAC---GGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 417
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420

Qy        418 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480

Qy        478 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540

Qy        538 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600

Qy        598 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCAGGCCGGCCCC 657
              ||||||||||||||||||||||||||||||||||||| ||||||||||||| ||||||||
Db        601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCC 660

Qy        658 GTGGCCCCCGGCCAGATGCGCGACCCCCGCGGCAGCGACATCGCCGGCGCCACCAGCACC 717
               | |||||||||||||||||||| |||||||||||||||||||||||| |||||||||||
Db        661 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 720

Qy        718 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCGTGCCCGTGGGCGACATCTAC 777
              ||||||||||||||||||||||||||||||||||||||| | ||||||||||||||||||
Db        721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780

Qy        778 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840

Qy        838 CTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 897
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900

Qy        898 ACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTG 957
              ||||||||||||||||||  ||||||||| |||||||||||||||||||| |||||||||
Db        901 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960

Qy        958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCACC 1017
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGC 1020

Qy       1018 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1077
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
```

RESULT 9
US-10-190-435-18

Sequence 18, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 3162
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmut_C

US-10-190-435-18

Query Match     84.5%;  Score 1275.2;  DB 15;  Length 3162;
Best Local Similarity   92.5%;  Pred. No. 2.6e-265;
Matches 1398;  Conservative  0;  Mismatches 78;  Indels 36;  Gaps  4;

RESULT 10
US-10-190-435-16

Sequence 16, Application US/10190435
Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 3462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagProtinaRTmut_C

US-10-190-435-16

Query Match　　84.5%;　Score 1275.2;　DB 15;　Length 3462;
Best Local Similarity　92.5%;　Pred. No. 2.6e-265;
Matches 1398;　Conservative　0;　Mismatches　78;　Indels　36;　Gaps　4;

Qy 1438 GACCGCGGAGCCCTACCGGACGCCCTGACGCCCTGTTCGTGGAGCGGCGCCC 1497
Db 1426 G----------------------ACCGGAGACCCTGACCAGCCTGAAGAGCGCCTGTTCGGCAACGACCCC 1473

Qy 1498 CTGAGCCAGTAA 1509
Db 1474 CTGAGCCAGAAA 1485

RESULT 11
US-10-190-435-19
; Sequence 19, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 4419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmutTatRevNef_C
US-10-190-435-19

Query Match 84.5%; Score 1275.2; DB 15; Length 4419;
Best Local Similarity 92.5%; Pred. No. 2.5e-265;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

Qy 1 ATGGGGCGCCCGCGCCCAGCCATCCTTGCGGGTGCGGGAGGAAGAGCTGGACAAGATTCCGC 60
Db 7 ATGGGGCGCCCGCGCCAGCCATCCTGGCGGCGGCAGCCATCTAC 66

Qy 61 CTGGCCCCCGGCCGCAAGAAGCACTACATGCTTGAAGCAAGTGCTGGAGGGCCATCCGCCGAG 120
Db 67 CTGGGCCCCGGGCGCAAGAAGTGCTACATGATGAAGCACCTTGTTGTGGGCCAGCCCGCGAG 126

Qy 121 CTGGAGGGGCTTCGCCCTGAACCCCGCGCCTGCTGGAGACCGCGAGGGGCTGCAAGCAGATC 180
Db 127 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 186

Qy 181 ATGAAGCAGCTGCACCCCGTGCACGCGGGCCACCGGGGCCGGGCGGCTGCGACAAC 240
Db 187 ATCCCCAGCCTGCACCCCGCCGCAGGCGCCAGGCGAGGAGCCTGAAGAGCCTGTTCAAC 246

Qy 241 ACCGTGGGCCACCTGTACTCCGTGCACGCGGAAAGATCGGAGTCCGCGGACACACAAGAGGAGGCC 300
Db 247 ACCGTGGGCCACCTGTACTCCGGCGAAAGATCGGAGTCCGGA----CCCCGACGCGCC 306

Qy 301 CTGACAAGATCGAGGAGGAGACAAGCTCCAGCAGAGGACCCAGCAGCAAGGAG 360
Db 307 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGCAGATCCAGGAGGCC 366

Qy 361 GCCGAC---GGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 417
Db 367 GCCGACAAGGGCCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACATCTAC 426

Qy 418 GTGCACCAGCCCATCAGCCCCGCACCCTGAAACGGCTGGGGTTGAAGGTGAATCGAGGCCAAGAG 477
Db 427 GTGCACCAGCCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 486

Qy 478 GCCTTCAGCCCCGAGGTGATCCCGCCCTTGACGGAGGCGGCGCCACCCCCCCCAG 537
Db 487 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 546

Qy 538 ... (alignment continues)
Db 538 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCGGCGCCATGCAGATCTGAAG 597

Qy 547 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 606
Db 547 GACTTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 606

Qy 598 GACACCATCAACGAAGAGGCCGGCCCCGAGTGGGACCGCCTGCACCCCGTTCAGCCGCCCC 657
Db 607 GACACCATCAACGAAGAGGCCGGCCGCGAGTGGGACCGCCTGCACCCCGTGCAGCCCCCC 666

Qy 658 GTGGCCCCCGGCCAGATGCGGGACCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 717
Db 667 ATCGGCCCCCGGCCAGATCCGCGAACCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 726

Qy 718 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCGTGCCCGTGGGCGACATCTAC 777
Db 727 CTCCAGGAGCAGGCACGGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 786

Qy 778 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCGCTGTCGATC 837
Db 787 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCTGTGAGCATC 846

Qy 838 CTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 897
Db 847 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 906

Qy 898 ACCCTGCGCGCCGAGCAGGCCACCCCAGACGTGAAGAACTGGATGACGGAGACCCTGCTG 957
Db 907 ACCCTGCGCGCCGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 966

Qy 958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCCTGGGCCCGGCCGCCACC 1017
Db 967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCCTGGGCCCCGCCGCCAGC 1026

Qy 1018 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCACAAGGGCCCCGCGCTG 1077
Db 1027 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCACAAGGGCCCCGCGCTG 1086

Qy 1078 CTGGCCGAGGCGCAGATGAGCCGCCAGGCCAACATCATGATGCAGAAGAGCAACTTC 1137
Db 1087 CTGGCCGAGGCGCAGATGAGCCGCCAGGCCAACAC-----AGCGTGATGATGCAGAGGAGCAACTTC 1143

Qy 1138 AAGGGCCCCCGCCGCGGCAAGCGTGAAGTGCTTCAACGTGGCAAGGAGGGCCACACCGCCAAG 1197
Db 1144 AAGGGCCCCCGCCGCGGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAG 1203

Qy 1198 AACTGCCGCGCCCCGCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCACCAGATG 1257
Db 1204 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCACCAGATG 1263

Qy 1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1317
Db 1264 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1323

Qy 1318 CGCCCCGGCAACTTCCTGCAGAACCGCCCGGAGCCCACCGCCCCACCGTGGCCCCCG 1377
Db 1324 CGCCCCGGCAACTTCCTGCAGAACCGC------CCCCGAGCCCACCGCC 1365

Qy 1378 CCCCCGCCGCAGAGCTTCGCGCTTCGGCGAGGAGACCACCCCGAGGCCCAAG 1437
Db 1366 CCCCCGCCGCGAGAGCTTCGGCTTCGGCGAGGAGACCACCCCGAGGCCCAAG 1425

Qy 1438 GACCGCGGAGCCCTACCGGACGCCCTGACGCCCTGTTCGTGGAGCGGCGCCC 1497
Db 1426 G------------ACCGGAGACCCTGACCAGCCTGAAGAGCGCCTGTTCGGCAACGACCCC 1473

Qy 1498 CTGAGCCAGTAA 1509
Db 1474 CTGAGCCAGAAA 1485

RESULT 12
US-10-190-305A-14

```
; Sequence 14, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: BARNETT, Susan
;  APPLICANT: LIAN, Ying
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
;  TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: 2302-18702 / 18702.002
;  CURRENT APPLICATION NUMBER: US/10/190,305A
;  CURRENT FILING DATE: 2002-07-05
;  NUMBER OF SEQ ID NOS: 93
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
;  LENGTH: 4419
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:
;  OTHER INFORMATION: GagRtmutatRevNef_C
US-10-190-305A-14

Query Match     84.5%;  Score 1275.2;  DB 15;  Length 4419;
Best Local Similarity   92.5%;  Pred. No. 2.5e-265;
Matches 1398;  Conservative   0;  Mismatches  78;  Indels  36;  Gaps  4;
```

```
Qy     1 ATGGGCGCCCGCGCGCCAGCATCCTGCGCGTGGCGGCGAGAAGTGGGAGAAGATCCGC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     7 ATGGGCGCCCGCGCGCCAGCATCCTGCGCGTGGCGGCGAGAAGTGGGAGAAGATCCGC 66

Qy    61 CTGCGCCCCGGCGGGCGGCGAAGAAGACTACATGCTGAAGCACCTCGGTGTGGGGCCAG 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67 CTGCGCCCCGGCGGGCGGCAAGAAGTCGCTACATAGATGAAGCACCTCGGTGTGGGGCCAG 126

Qy   121 CTGGGAGGGGCTTCGCCCTGGAACCCCCGGGGCTGCTGTGGAGACCCGGGCTGCAAGCAGATC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127 CTGGGAAAGTCGCCCTGAACCCCCGGGGCCTGCTGTGGAGACCCGGGCTGCAAGCAGATC 186

Qy   181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACACCGGCCACGGAGCTGCCAGCCTGTTCAAC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187 ATCCGCGACTCACCCCGCCTGCAGACACCGGCCACCGAGCTGAAGAACCTGTTCAAC 246

Qy   241 ACCGTGGCCACCCTGTACTGCGTGCACCCGGGGCATCAGGTCGGCGGGTACAAGGAGGCC 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247 ACCGTGGCCACCCTGTACTGCGTGCACCAGAAGATCGAGGTCCGGGACACCAAGGAGGCC 306

Qy   301 CTGGACAAGATCGAGGAGGAGGAGAACAAGTCCCAGCAGCAGCCCAAGGAG 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307 CTGGACAAGATCGAGGAGGAGAACAAGTGCCAGCAGAAGATGCAGCAGGCC 366

Qy   361 GCCGAC---GGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 417
         |||||||   |||||||||||||||||||||||||||||||||||
Db   367 GCCGACAAGAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGGCCAGATG 426

Qy   418 GTGCACCAGGCCATCAGCCCCTGAACGCCCGGGGGCGCCACCCCCAG 477
         |||||||||||||||||||||||||||||||||||||||||||||
Db   427 GTGCACCAGGCCATCAGCCCCCGGAGCGTGAATCAGGGGCCACCCCCCAG 486

Qy   478 GCCTTCAGCGCCGAGGGTGATCCCCATGGTTCACGGCCCAGGCGGGA 537
         |||||||||||||||||||||||||||||||||||||||||||||
Db   487 GCCTTCAGCGCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCAG 546

Qy   538 GACCTGAACACGATGTTGAACACCGTGGGCGGGGCACCGGCCACCCAGGCCATCAGCCCC 597
         |||||||||||||||||||||||||||||||||||||||
Db   547 GACCTGAACACGATGTTGAACACCGTGGGCGGGGCACCGGCCACCCAGGGGATGCTGAAG 606

Qy   598 GACACCATCAACGAGGAGGCCGCCGGGTGGGACGGCCCCCG 657
         |||||||||||||||||||||||||||||||||||||||||
Db   607 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGCTGCACCCGGTGCACGCCGGCCCC 666

Qy   658 GTGGCCCCCCGGCCCCGAGGATCGCGGACTCGCCCGGGCACCCACCAGCACC 717
         |||||||||||||||||||||||||||||||||||||||||
Db   667 ATCGCCCCCGGCCAGATGCGCGAGCCGAGGGGCTCCGACATCGCCGGCAACCACGAGCACC 726

Qy   718 CTGCAGGGACAGATCGCCTGGATGACCAGCAACCCCCCCGTGCCCGTGGGCGACATCTAC 777
         |||||||||||||||||||||||||||||||||||||||||
Db   727 CTGCAGGGACAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 786

Qy   778 AAGCGGTGGATCATCCTGGGCCTGAACAAAGTCGTGCGGATGTACAGCCCCGTGAGCATC 837
         |||||||||||||||||||||||||||||||||||||||||
Db   787 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 846

Qy   838 CTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 897
         |||||||||||||||||||||||||||||||||||||||||
Db   847 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 906

Qy   898 ACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTG 957
         |||||||||||||||||||||||||||||||||||||||||
Db   907 ACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTG 966

Qy   958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGCGCCACC 1017
         |||||||||||||||||||||||||||||||||||||||||
Db   967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGCGCCAGC 1026

Qy  1018 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1077
         |||||||||||||||||||||||||||||||||||||||||
Db  1027 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1086

Qy  1078 CTGGCCGAGGCGGATGAGCCAGGGCCAACAGCGGTGAACATCATGATGCAGAAGAGCAACTTC 1137
         |||||||||||||||||||||||||||||||||||||||
Db  1087 CTGGCCGAGGCGATGAGCCAGGCCAACAGC---AGCGTGATGATGCAGAAGAGCAACTTC 1143

Qy  1138 AAGGGCCCCCGGCGCAACGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGC 1197
         |||||||||||||||||||||||||||||||||||||||
Db  1144 AAGGGCCCCCGGCGCAACGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGC 1203

Qy  1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257
         |||||||||||||||||||||||||||||||||||||||
Db  1204 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1263

Qy  1258 AAGGACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1317
         |||||||||||||||||||||||||||||||||||||||
Db  1264 AAGGACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1323

Qy  1318 CGCCCCGGCAACTTCCTGCAGAACCGCCCAGAGCCCACCGCCCCACCCGCC 1377
         |||||||||||||||||||||||||||||||||||||||
Db  1324 CGCCCCGGCAACTTCCTGCAGAGCCG---------CCCGAGCCCACCGCC 1365

Qy  1378 CCCCCGCCCGAGAGCTTCGGCTTCGGAGGAGACCACCCCCCAGAGCCCAAG 1437
         |||||||||||||||||||||||||||||||||||||||
Db  1366 CCCCCGCCCGGAGCTTCGGCTTCGAGGAGACCACCCCCGCCCCAGAGGAGCAG 1425

Qy  1438 GACCGGCAGCCCTACCGCGCCCTGACCCGCCGCCAGCCGTTCGCGCAGCGCGCCC 1497
         |||||||||||||||||||||||||||||||
Db  1426 G------------ACCGCGAGACCCCGACCCGCCGCCAGCCGTTCGGCAACGACCCCC 1473

Qy  1498 CTGAGCCAGTAA 1509
         ||||||||||||
Db  1474 CTGAGCCAGAAA 1485
```

RESULT 13
US-10-190-435-35
; Sequence 35, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: BARNETT, Susan W.
;  APPLICANT: LIAN, Ying
;  APPLICANT: ENGELBRECHT, Susan
;  APPLICANT: VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP10133.003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE: 2002-12-30

```
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: Patentin Ver. 2.0
;
;  SEQ ID NO 35
;  LENGTH: 4483
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:
;  OTHER INFORMATION: gp160mod.TV1.dVIV2-gagmod.BW965
US-10-190-435-35

Query Match      84.5%;   Score 1275.2;  DB 15;  Length 4483;
Best Local Similarity   92.5%;   Pred. No 2.5e-265;
Matches 1398;   Conservative  0;   Mismatches 78;   Indels  36;   Gaps  4;
```

```
Qy      1  ATGGGCGGGGCCCCCGGGGGCCCCGGCCAGCCATCCTGCGCGGGCCGGAGGAAGCTGGACAAGTGGGAGAAAGATCCGC   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3005  ATGGGCGGGGCCCCCGGGGGCCCCGGCCAGCATCCTGCGCGGGCCGGAGGAAGCTGGACAAGTGGGAGAAAGATCCGC  3064

Qy     61  CTGCGCCCCGGCGGGCGGGGCCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG              120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3065  CTGCGCCCCGGCGGGCGGGGCCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG  3124

Qy    121  CTGGAGAGGGCTTCGCCCCTGAACCCGGCCCGCCTGCTGGGACCGGGGGGCTGCAAGCAGATC                180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3125  CTGGAGAGAGTTCGCCCCTGAACCCGGCCCGCCTGCTGGGAGACCCAGGGGCTGCAAGCAGATC  3184

Qy    181  ATGAAGCAGCTGCAGCCCCTGCAGCCGACCGGCCAGCCGCCCCCGCCGCCGCCGCCTGTGCAAC               240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3185  ATCGCGCCAGCTGCACCCCCCGCTGCAGACCGGCCAGGACCTGCAGGAGCCTGTTCAAC  3244

Qy    241  ACCGTGGCCCACCCTGTTACTGGGTGCACGCCCGGCATCGAAGGTCGCGGACAAGAGGCC                 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3245  ACCGTGGCCCACCCTGTTACTGCGTGCAGCGAGAAGATGCGGTCGCGGACAAGAGGGCC  3304

Qy    301  CTGGACAAGATCGAGGAGGAGACAAGTCCCAGCAGACAGCCCAGGATGCCCAAGGAG                   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3305  CTGGACAAGATCGAGGAGGAGACAACAAGTCCAGCAGACAGCCCAGGATGCCCAAGGCC  3364

Qy    361  GCCGGAC---GGCAAGGTGAGCCAGAAACTACCCCCATCGTGTGCAGAGACTTGCAGGGCCAGATC             417
           |||||||   |||||||||||||||||||||||||||||||||||||||||
Db   3365  GCCGGACAAGGGGGACCCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCAGGGCCAGATC  3424

Qy    418  GTGCACCAGGCCATCAGCCCCGAGTGATCCCGCCCCGCACCCTGAACGCCCATCGGGGTGAAGTGATCGAGGAGAG    477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3425  GTGCACCAGGCCCATCAGCCCCGCCACCCTGAACCGCCATGGGGGTGGGTGAAGTGATCGAGGAGAAG  3484

Qy    478  GCCTTCAGCCCCGAGGTGATCCCATGTTCACGCCCCTGACGCGAGGGCGCCACCCCCCCAG                537
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3485  GCCTTCAGCCCCGAGGTGATCCCATGTTCACCGGATGTTCCACGCCCTGACGGCCACCCCCCCCAG  3544

Qy    538  GACCTGAACACGACGATGTTGAACACCCTGGGCGGCCGCACCACCCTGCAGGCGCTGATGCTGAAG           597
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3545  GACCTGAACACGATGTTGAACACCCTGGGCGGCCGCACCACCCTGCAGGCGCTGATGCTGAAG  3604

Qy    598  GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCGCCACCCGGTGCCAGGCCGCCCGGCCCGGGCCCC          657
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3605  GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCGGTGCCAGGCCGCCCGGGCCCC  3664

Qy    658  GTGGCCCCCGGCCCAGATGCGCGCACATGCGCAGGCCACCAGCACC                      717
           |||||||||||||||||||||||||||||||||||||||||||||
Db   3665  ATCGCCCCGGCCCCAGATGGCGAGCCGCGGCGACCCCGGCGCCACATCGCCAGCACC  3724

Qy    718  CTGCAGGAGCAGATCGGCTGGATGACCAGCAACCCCGTGCCCGTGGGCGACATCTAC                  777
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3725  CTGCAGGAGCAGATCGGCTGGATGACCAGCAACCCCGTGGGCGACATCTAC  3784

Qy    778  AAGCGGGTGGGATCATCCTGGCGCCTGAACACAGATGCGTGGGATGTGTACAAGCGCCCGGACAGCATC          837
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3785  AAGCGGTGGATCATCCTGGCGCCTGAACACAGATGCGTGGGATGTGTACAAGCGCCCGGACAGCATC  3844

Qy    838  CTGGACATCCGCCAGGGACCCCAAGGAGCCCTTCGGCGACTACGTGGACCGGCCTTCTTCAAG            897
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3845  CTGGACATCCAAGCAGCAGGGCCCCCAAGGAGCCCCAAGGAGGAGCCCTTCGGCGACTACGTGGACCGGCCTTCTTCAAG  3904
```

```
Db   3905  ACCCTGCGCGCCGGCCAGATGAGCCAGGCGAGTGAACCGGAGCACCCGGAGGTGAACACGGAGAACTGGATGACCGAGACCCTGCTG  3964

Qy    958  GTGCAGGACGGCCAACCCCGGACTTGCAGAGACCATCCTGCGCGCCTTCTTCGGCCGCCGCCCGCCCGGGGGGGCCCACC  1017
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3965  GTGCAGGACGGCCAACCCCGGACTTGCAGAGACCATCCTGCGCGCCTTCTTCGGCCGCCGCCCGCCCGGGGGGGCCCACC  4024

Qy   1018  CTGGAGGAGGATGAGTGACCGGCTGCCAGGGCGTGGGGGTGGGGCGCCCCACCAAGGCCCGGGTG  1077
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4025  CTGGAGGAGGATGATGAGTGACCGGCTTGCCAGGGCGTGGGGGTGGGGCGCCCCACCAAGGCCCGGGTG  4084

Qy   1078  CTGGCCGGAGGCGGATGAGCGCAAGCGTGAACATGCATGATGCAGAGAGAGCAACTTC  1137
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   4085  CTGGCCGAGGCGGATGAGCGCAAGCGTGAACATGCATGATGCAGAGAGAGCAACTTC  4141

Qy   1138  AAGGGCCCCCCCCGGGGCCCGGGGGCCCATCCTCAACGTGCTTCAACTGCGCGACACCGGAC---AGCGTGATGATCGCAGAGAGAGAGCAACACC  1197
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db   4142  AAGGGGCCCCCCGGGGCCCCATCCTCAACTGCTTCAACTGCGGCAAGGAGGGGCACATCGCCAGCAGCAACTGCAGGGCCCCAAGGGC  4201

Qy   1198  AACTGCCGGGCCCCCGCCCCGGGCCGCGAAGAAGGGTGCTGGAAGTGCGGCCAAGGCCGCCAAGGAGGGGGCCCATCGCCCCGC  1257
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4202  AACTGCCGGGCCCCCGCCCCGGGCCGCGAAGAAGGGTGCTGGAAGTGCGGCCAAGGAGGGGGCCCATCGCCCCGC  4261

Qy   1258  AAGGACTGCACCGAGCGCCAGGCCAACCTTCCTCGGGCCAAGTGTCTGGCCCAGCCAGTCAGGGGC  1317
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4262  AAGGACTGCACCGAGCGCCAGGCCAACTTCCTCGGGCCAACTCGGCCCAGCCGCACAGAGGGC  4321

Qy   1318  CGGCCCCCGGCCAACTTCCTCCGAGAACCCGAGCGGCGGCCCGCGGCCCCCACCGGTTGCCCCACCGCC  1377
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4322  CGCCCCCGGCCAACTTCCTCCGAGACCG-------------CCCCGAGCCCCACCGGC  4363

Qy   1378  CCCCCGCGCCCGAGAGCTTCCGCTTCGAGGAGGACAACCCCCCCCCAAGCAGAGGAGCC  1437
           |||||||||||||||||||||||||||||||||||||||||
Db   4364  CCCCCGCGCCCGAGAGCTTCCGCTTCGAGGAGGAGCCACCCCCGGCCGCCAAGGAGCAAG  4423

Qy   1438  GACCGGCGAGACCCTTACCGCCGAGCCCCTGGCGCCCCCC                           1497
           |||||||||||||||||||||||||||||
Db   4424  G--------ACCGCGAGACCCTGACCAGCCTGAAGAGCCCTGTTCGGCCACGACCCC  4471

Qy   1498  CTGAGCCAGTAA  1509
           ||||||||||||
Db   4472  CTGAGCCAATAA  4483
```

Query Match   84.5%;   Score 1275.2;   DB 15;   Length 4606;
Best Local Similarity   92.5%;   Pred. No. 2.5e-265;
Matches 1398;   Conservative   0;   Mismatches   78;   Indels   36;   Gaps   4;

*(DNA sequence alignment between Query (Qy) and Database (Db) sequences, positions 1–1077 / 3128–4207 shown as paired alignment rows.)*

RESULT 15
US-10-190-435-36
; Sequence 36, Application US/10190435
; Publication No. US2003014324BA1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 36
; LENGTH: 4615
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TV1.dV2-gagmod.BW965
US-10-190-435-36

Query Match   84.5%;   Score 1275.2;   DB 15;   Length 4615;
Best Local Similarity   92.5%;   Pred. No. 2.5e-265;
Matches 1398;   Conservative   0;   Mismatches   78;   Indels   36;   Gaps   4;

*(DNA sequence alignment between Query (Qy) and Database (Db) sequences follows.)*

(Nucleotide sequence alignment — Qy (query) / Db (database) paired rows with position numbers. The aligned sequence characters are not individually reproduced here.)

QM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:          June 1, 2004, 10:26:25 ; Search time 612 Seconds
                 (without alignments)
                 10474.724 Million cell updates/sec

Title:           US-09-475-704A-4
Perfect score:   1509
Sequence:        1 atgggcgccgccaagcat..........gcggcccctgagccagtaa 1509
Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched:        3373863 seqs, 212409041 residues
Total number of hits satisfying chosen parameters:     6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq 29Jan04:*
                 1: geneseqn1980s:*
                 2: geneseqn1990s:*
                 3: geneseqn2000s:*
                 4: geneseqn2001as:*
                 5: geneseqn2001bs:*
                 6: geneseqn2002s:*
                 7: geneseqn2003as:*
                 8: geneseqn2003bs:*
                 9: geneseqn2003cs:*
                10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | Match DB ID | Description |
|---|---|---|---|---|---|
| 1 | 1509 | 100.0 | 1509 | 3 | AAA51610 | Aaa51610 HIV synth |
| 2 | 1509 | 100.0 | 1509 | 6 | AAL44549 | Aal44549 HIV-1 p55 |
| 3 | 1509 | 100.0 | 1509 | 6 | ABL39954 | Abl39954 Synthetic |
| 4 | 1491.4 | 98.8 | 1509 | 6 | ABL39958 | Abl39958 Synthetic |
| 5 | 1489.8 | 98.7 | 1509 | 3 | AAA51626 | Aaa51626 HIV codon |
| 6 | 1489.8 | 98.7 | 1494 | 6 | AAL44553 | Aal44553 HIV-1 p55 |
| 7 | 1297.6 | 86.0 | 1494 | 6 | ABL39972 | Abl39972 Synthetic |
| 8 | 1294.8 | 85.8 | 1491 | 5 | ABL40020 | Abl40020 Synthetic |
| 9 | 1296.8 | 84.6 | 1479 | 5 | AAA51609 | Aaa51609 HIV synth |
| 10 | 1276.8 | 84.6 | 1479 | 6 | AAL44548 | Aal44548 HIV-1 p55 |
| 11 | 1276.8 | 84.6 | 1479 | 6 | ABL39953 | Abl39953 Synthetic |
| 12 | 1276.2 | 84.6 | 3162 | 9 | ADC13239 | Adc13239 DNA of HI |
| 13 | 1275.2 | 84.5 | 3462 | 9 | ADC13237 | Adc13237 DNA of HI |
| 14 | 1275.2 | 84.5 | 4419 | 7 | ACA03523 | Aca03523 Synthetic |
| 15 | 1275.2 | 84.5 | 4419 | 7 | ADC13240 | Adc13240 DNA of HI |
| 16 | 1275.2 | 84.5 | 4615 | 9 | ADC13257 | Adc13257 DNA of HI |
| 17 | 1275.2 | 84.5 | 4702 | 9 | ADC13259 | Adc13259 DNA of HI |
| 18 | 1275.2 | 84.5 | 4716 | 7 | ACA03522 | Aca03522 Synthetic |
| 19 | 1275.2 | 84.5 | 4716 | 9 | ADC13238 | Adc13238 DNA of HI |
| 20 | 1273.6 | 84.4 | 2742 | 7 | ACA03524 | Aca03524 Synthetic |
| 21 | 1273.6 | 84.4 | 2742 | 9 | ADC13241 | Adc13241 DNA of HI |
| 22 | 1273.6 | 84.4 | 3930 | 9 | ADC13230 | Adc13230 DNA of HI |
| 23 | 1273.6 | 84.4 | 3930 | 9 | ADC13231 | Adc13231 DNA of HI |
| 24 | 1273.6 | 84.4 | 3930 | 9 | ADC13232 | Adc13232 DNA of HI |
| 25 | 1273.6 | 84.4 | 5145 | 7 | ACA03521 | Aca03521 Synthetic |
| 26 | 1273.6 | 84.4 | 5145 | 9 | ADC13233 | Adc13233 DNA of HI |
| 27 | 1272.6 | 84.3 | 4713 | 7 | ACA03592 | Aca03592 Synthetic |
| 28 | 1272.6 | 84.3 | 4713 | 9 | ADC13280 | Adc13280 DNA of HI |
| 29 | 1271.6 | 84.3 | 5184 | 7 | ACA03591 | Aca03591 Synthetic |
| 30 | 1271.6 | 84.3 | 5184 | 9 | ADC13279 | Adc13279 DNA of HI |
| 31 | 1270.2 | 84.2 | 2742 | 7 | ACA03590 | Aca03590 Synthetic |
| 32 | 1270.2 | 84.2 | 2742 | 9 | ADC13278 | Adc13278 DNA of HI |
| 33 | 1260.8 | 83.6 | 1479 | 9 | AAA51625 | Aaa51625 HIV synth |
| 34 | 1260.8 | 83.6 | 1479 | 6 | ABL39957 | Abl39957 Synthetic |
| 35 | 1259.2 | 83.4 | 1479 | 6 | AAL44552 | Aal44552 HIV-1 p55. |
| 36 | 1234.4 | 81.8 | 4546 | 9 | AAH20868 | Aah20868 HIV-1 sub |
| 37 | 1202.2 | 79.7 | 4288 | 4 | AAH20868 | Aah20868 HIV-1 sub |
| 38 | 1201 | 79.6 | 1515 | 3 | AAA70412 | Aaa70412 HIV bicis |
| 39 | 1201 | 79.6 | 4608 | 3 | AAA70412 | Aaa70412 HIV bicis |
| 40 | 1201 | 79.6 | 4689 | 3 | AAA70473 | Aaa70473 HIV bicis |
| 41 | 1201 | 79.6 | 4766 | 3 | AAA70470 | Aaa70470 HIV bicis |
| 42 | 1201 | 79.6 | 2799 | 3 | AAA70471 | Aaa70471 HIV bicis |
| 43 | 1199.4 | 79.5 | 2799 | 7 | ACA03520 | Aca03520 Synthetic |
| 44 | 1199.4 | 79.5 | 2799 | 7 | ACC78493 | Acc78493 HIV GagTa |
| 45 | 1199.4 | 79.5 | 3205 | 7 | ACC78492 | Acc78492 HIV GagRT |

ALIGNMENTS

RESULT 1
AAA51610
ID   AAA51610 standard; DNA; 1509 BP.
XX
AC   AAA51610;
XX
DT   31-OCT-2000  (first entry)
XX
DE   HIV synthetic Gag polynucleotide.
XX
KW   Gag; expression cassette; antigenic; type C; HIV; synthetic;
KW   DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS   Human immunodeficiency virus; type C strain AF110967.
OS   Synthetic.
XX
PN   WO200039304-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;  99WO-US031273.
XX
PR   31-DEC-1998;  98US-0114495P.
PR   01-SEP-1999;  99US-0152195P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S,  Zur Megede J;
XX
DR   WPI; 2000-452401/39.
XX
PT   Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env
PT   polypeptide and the polypeptide useful for immunizing a mammal especially
PT   human against HIV.
XX
PS   Claim 2; Page 93; 113pp; English.
XX
CC   Expression cassettes comprising a polynucleotide encoding antigenic type
CC   C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful
CC   in DNA immunization, generation of packaging cell lines and production of
CC   Gag- and/or Env-containing proteins. Synthetic Env and Gag expression
CC   cassettes exhibit increased potency for induction of cytotoxic T-
CC   lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-
CC   assemble into non-infectious virus-like particles which are used as a
CC   matrix for the proper presentation of an antigen entrapped or associated
CC   to the immune system of the host

```
XX
SQ   Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other;
Query Match         100.0%;  Score 1509;  DB 3;  Length 1509;
Best Local Similarity   100.0%;  Pred. No. 3.8e-177;
Matches 1509;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy       1 ATGGGGCCCGGGCCCCAGCATCCTCGCGGGCGGGGCGAGAAGTGGGAGAGAGATCCGC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGGGCCCGGGCCCCAGCATCCTCGCGGGCGGGGCGAGAAGTGGGAGAGAGATCCGC 60

Qy      61 CTGCGCCCCGGGGGGGCGGCCAAGAAGCACTACATGCTGAAGCACGTGGGACAAGTGGGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTGCGCCCCGGGGGGGCGGCCAAGAAGCACTACATGCTGAAGCACGTGGGACAAGTGGGAG 120

Qy     121 CTGGGAGGGGCTTCGCCCTTGGAAGAACCCCGGGGCGACTGTGAAGCGCGCGGAGAAGATC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTGGGAGGGGCTTCGCCCTTGGAAGAACCCCGGGGCGACTGTGAAGCGCGCGGAGAAGATC 180

Qy     181 ATGAAGCAGCTGCAGCCGGCCCTGCAGACCGGCAGCGAGGAGCTGCGGAGCCTGTACAAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATGAAGCAGCTGCAGCCGGCCCTGCAGACCGGCAGCGAGGAGCTGCGGAGCCTGTACAAC 240

Qy     241 ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTCAAGGACACCAAGGAGGCC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTCAAGGACACCAAGGAGGCC 300

Qy     301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGCAGCGGAG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGCAGCGGAG 360

Qy     361 GCGGACAAGGTGAGCCAAGAACTACCCCATGGTGCAGAACCTGCAGGGCCCAGATGGTG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GCGGACAAGGTGAGCCAAGAACTACCCCATGGTGCAGAACCTGCAGGGCCCAGATGGTG 420

Qy     421 CACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATGGAGGAAGAGGCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATGGAGGAAGAGGCC 480

Qy     481 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCGCTGAGCGAGGGGGCCACCCCCAGGAC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCGCTGAGCGAGGGGGCCACCCCCAGGAC 540

Qy     541 CTGAACACGATGTTGAACACCGTGGGGGCCCACCAGGCGGCCCATGCAGATGCTGAAGGAC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CTGAACACGATGTTGAACACCGTGGGGGCCCACCAGGCGGCCCATGCAGATGCTGAAGGAC 600

Qy     601 ACCATCAACGAGGAGGCGCCGCCAGTGGGACCCGGCCTGCAGGGCCGGCCCCCCGTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ACCATCAACGAGGAGGCGCCGCCAGTGGGACCCGGCCTGCAGGGCCGGCCCCCCGTG 660

Qy     661 GCCCCCGGCCAGATGCGGGAACCCCGCGGCAGCGACATGCGCACCAGCACCCTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GCCCCCGGCCAGATGCGGGAACCCCGCGGCAGCGACATGCGCACCAGCACCCTG 720

Qy     721 CAGGAGCAGATCGCCTGGATGACCAGCAACCCACCCATCCCCGTGGGCGACATCTACAAG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CAGGAGCAGATCGCCTGGATGACCAGCAACCCACCCATCCCCGTGGGCGACATCTACAAG 780

Qy     781 CGGTGGATCATCCTGGGCCTGAACAAGATGTGCGGAGTACAGCCCGGCGGCATCCTG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CGGTGGATCATCCTGGGCCTGAACAAGATGTGCGGAGTACAGCCCGGCGGCATCCTG 840

Qy     841 GACATCCGGCCAGGGCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAGACC 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GACATCCGGCCAGGGCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAGACC 900

Qy     901 CTGCGCGCCGAGCAGGCCACCCAGGACGTTGAAGAACTGGACATGACCCGAGGACCCGTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTGCGCGCCGAGCAGGCCACCCAGGACGTTGAAGAACTGGACATGACCCGAGGACCCGTG 960

Qy     961 CAGAACGCCAACCCGGACTGCAAGACCATCCTGCGCGCGCTCGGCCCCGGGCGCCACCCTG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 CAGAACGCCAACCCGGACTGCAAGACCATCCTGCGCGCGCTCGGCCCCGGGCGCCACCCTG 1020

Qy    1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1080

Qy    1081 GCCCGGCGGCATGAGCCAGGCCAACAGCGTGAAACATCATGTCATGCAGAAGAGCCAACTTCAAG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GCCCGGCGGCATGAGCCAGGCCAACAGCGTGAAACATCATGTCATGCAGAAGAGCCAACTTCAAG 1140

Qy    1141 GGCCCCCCGGCGCCAAGGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAAC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GGCCCCCCGGCGCCAAGGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAAC 1200

Qy    1201 TGCCGCGCCCCCAGGAGGCGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TGCCGCGCCCCCAGGAGGCGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260

Qy    1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCTACAAGGGCCGC 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCTACAAGGGCCGC 1320

Qy    1321 CCCGGCAACTTCCTGCAGAACCGCCCCGAGCCCACCGCACCGTGCCCACCGTGCCCC 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CCCGGCAACTTCCTGCAGAACCGCCCCGAGCCCACCGCACCGTGCCCACCGTGCCCC 1380

Qy    1381 CCCCGGCCGAGGCTTCGGCTTCGGGGAGGACCACCCCCCCCCGCCCCCCAGCCCCAAGGAC 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 CCCCGGCCGAGGCTTCGGCTTCGGGGAGGACCACCCCCCCCCGCCCCCCAGCCCCAAGGAC 1440

Qy    1441 CGCCAGACGCCTACCGCGAGCCCCCTGTTCGGGCCAGCCCCCCTG 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 CGCCAGACGCCTACCGCGAGCCCCCTGTTCGGGCCAGCCCCCCTG 1500

Qy    1501 AGCCAGTAA 1509
           |||||||||
Db    1501 AGCCAGTAA 1509
```

RESULT 2
AAL44549
ID   AAL44549 standard; DNA; 1509 BP.
XX
AC   AAL44549;
XX
DT   29-AUG-2003  (revised)
DT   08-NOV-2002  (first entry)
XX
DE   HIV-1 p55gag polypeptide coding sequence 2.
XX
OS   Human immunodeficiency virus 1.
XX
PN   WO200226209-A2.
XX
PD   04-APR-2002.
XX
PF   28-SEP-2001; 2001WO-US030540.
XX
PR   28-SEP-2000; 2000US-02336105P.
PR   30-AUG-2001; 2001US-0315905P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;
PI   Ulmer J, Dubensky TW;
XX
DR   WPI; 2002-519084/55.

KW   HIV; ds; vaccine; gene; immune response; microparticle;
KW   adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
KW   polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
KW   submicron emulsion; viral infection; bacterial infection;
KW   parasitic infection; HIV-1 p55gag polypeptide.

XX
PT A microparticle to which a biologically active macromolecule is adsorbed,
PT for use as a vaccine composition to treat viral, bacterial or parasitic
PT infections, comprises a polymer microparticle, a detergent and a
PT submicron emulsion.
XX
PS
XX
PT Claim 72; Fig 2; 100pp; English.
XX
CC The invention relates to a method of raising an immune response in a host
CC animal. The method of the invention comprises administering a
CC microparticle that has an adsorbent surface to which a first biologically
CC active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC a polyanhydride, a polycyanoacrylate, a detergent, and submicron emulsion. The method/
CC microparticle of the invention is useful for immunising a host animal
CC against viral, bacterial or parasitic infections. The present DNA
CC sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other;

Query Match   100.0%;   Score 1509;   DB 6;   Length 1509;
Best Local Similarity   100.0%;   Pred. No. 3.8e-177;
Matches 1509;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 ATGGGCGCGCCCGCCAGCATCCTGCCGGGCGGACAGTGGGAAGAAGATCCGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGGCGCGCCCGCCAGCATCCTGCCGGGCGGACAGTGGGAAGAAGATCCGC 60

Qy     61 CTGCGCCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCGCCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 120

Qy    121 GAGGAGGGGCTTCGCCCTGAACCCGGGCCTGCTGGAGACCAGCGAGGGCCTGCAAGCAGATC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGGAGGGGCTTCGCCCTGAACCCGGGCCTGCTGGAGACCAGCGAGGGCCTGCAAGCAGATC 180

Qy    181 ATGAAGCAGCTGCAGCCCGCCCAGCAGACCGGACCCGGCCCACCGCGGCCAGCCTGTACAAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATGAAGCAGCTGCAGCCCGCCCAGCAGACCGGACCCGGCCCACCGCGGCCAGCCTGTACAAC 240

Qy    241 ACCGTGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTCCGCGCAGCACCAAGGAGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACCGTGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTCCGCGCAGCACCAAGGAGGCC 300

Qy    301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAG 360

Qy    361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATGCAGCTGAACGGCCAGATGGTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATGCAGCTGAACGGCCAGATGGTG 420

Qy    421 CACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCC 480

Qy    481 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCCCACCCCCAGGAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCCCACCCCCAGGAC 540

Qy    541 CTGAACACGATGTTGAACACCGGTGGGCCACCGGCCCCCGAGTGGGACCATGCTGAAGGAC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTGAACACGATGTTGAACACCGGTGGGCCACCGGCCCCCGAGTGGGACCATGCTGAAGGAC 600

Qy    601 ACCATCAAGGAGGAGGGCCCCGGAGTGGACCACGGCGGCCTGCACCCCGTGCACGCCCCGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACCATCAAGGAGGAGGGCCCCGGAGTGGACCACGGCGGCCTGCACCCCGTGCACGCCCCGTG 660

Qy    661 GCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACCCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACCCTG 720

Qy    721 CAGGAGCAGCGAATGGCCTGGATGACCAGCAACCCCCCGTGCCCGTGGGGGACCATCTACAG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CAGGAGCAGCGAATGGCCTGGATGACCAGCAACCCCCCGTGCCCGTGGGGGACCATCTACAG 780

Qy    781 CGGTGGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCGTGAGGCATCCTG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CGGTGGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCGTGAGGCATCCTG 840

Qy    841 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAGACC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAGACC 900

Qy    901 CTCCGGGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGGACCCTGCTGGTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTCCGGGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGGACCCTGCTGGTG 960

Qy    961 CAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCTCTCGGGCCCCCACCCTG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCTCTCGGGCCCCCACCCTG 1020

Qy   1021 GAGGAGGATGATGACCGCGCCCTGGGCGGGGCGGCCCACAAGGCCCGCGTGCTG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GAGGAGGATGATGACCGCGCCCTGGGCGGGGCGGCCCACAAGGCCCGCGTGCTG 1080

Qy   1081 GCCGAGGGCCATGAGCCAGCCCAACAGCGTGAACATCATGATGCAAGAGAGCAACTTCAAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GCCGAGGGCCATGAGCCAGCCCAACAGCGTGAACATCATGATGCAAGAGAGCAACTTCAAG 1140

Qy   1141 GGCCCCCGGCGCAAGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAAC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GGCCCCCGGCGCAAGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAAC 1200

Qy   1201 TGCCGCGCCCCCCGGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TGCCGCGCCCCCCGGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260

Qy   1261 GACTGCACCGAGGGCGGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GACTGCACCGAGGGCGGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCC 1320

Qy   1321 CCCGGCCAACTTCCTGCAGAACCGCAGCGAGCCCACCGCCCCCCCCCACCGTGCCCAAGGAC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CCCGGCCAACTTCCTGCAGAACCGCAGCGAGCCCACCGCCCCCCCCCACCGTGCCCAAGGAC 1380

Qy   1381 CCGGCCCAGAGGCTTCCGCGTTCGAGGAGACCACCCCCGCCCCCAAGCAGGAGCCCAAGGAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CCGGCCCAGAGGCTTCCGCGTTCGAGGAGACCACCCCCGCCCCCAAGCAGGAGCCCAAGGAC 1440

Qy   1441 CGGGAGCCCTACCGGCAGGCCTGTGTTCGGGCCCCCTG 1500
          ||||||||||||||||||||||||||||||||||||||||
Db   1441 CGGGAGCCCTACCGGCAGGCCTGTGTTCGGGCCCCCTG 1500

Qy   1501 AGCCAGTAA 1509
          |||||||||
Db   1501 AGCCAGTAA 1509
```

RESULT 3
ID   ABL39954 standard; DNA; 1509 BP.
AC   ABL39954;
XX
DT   15-MAY-2002   (first entry)
XX
DE   Synthetic Gag polynucleotide sequence SEQ ID NO:4.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.

OS   Synthetic.

XX

PN   WO20204493-A2.

XX

PD   17-JAN-2002.

XX

PF   05-JUL-2001; 2001WO-US021241.

PR   05-JUL-2000; 2000US-00610313.

XX

PA   (CHIR ) CHIRON CORP.

PA   (UYST-) UNIV STELLENBOSCH.

XX

PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg BJ;

XX

DR   WPI; 2002-154920/20.

XX

PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.

XX

PS   Example 1; Fig 2; 233pp; English.

XX

CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (II) have immunostimulant activity and can be used in applications
CC   including DNA immunisation, generation of packaging cell lines, and
CC   production of HIV
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention.

XX

SQ   Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other;

Query Match    100.0%;   Score 1509;   DB 6;   Length 1509;
Best Local Similarity   100.0%;
Matches 1509;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy          1 ATGGGCGCCCCCGGCGCCGCGCCATCCTGCGCGGGCGGGGAGCGGAGGAAGATCCGC   60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCGCCCCCAGCATCCTGCGCGGGCGGGGAGCGGAGGAAGATCCGC   60

Qy         61 CTGCGCCCCGGCGGGAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCGCCCCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG  120

Qy        121 CTGGAGAGGCTTCGCCCTGAAACCCCGGCCCTGCTGCAGGGCGCGCCCAAGAGCAGATC  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGAGAGGCTTCGCCCTGAAACCCCGGCCCTGCTGGGACCCTGCAAGCACATC  180

Qy        181 ATGAAGCAGCTTCAGCCCCGCCGCCACCGGCCACCGGAGGTGCCCAGCCTGTACAAC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGAAGCAGCTTCAGCCCCGCCGCGACCCGCGACCACCAGAGGAGGCC  240

Qy        241 ACCGTGGCCCACCCTGTACTGCGGTGCACCCGGGCATCGAGCTCCCGCGACACCCAAGGAGGCC  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACGGTGGCCCACCCTGTACTGCGGTGCATCGAGCTCCCGCGACACCCAAGGAG  300

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAC  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAC  360

Qy        361 GCCGGACGGCGAAGGTGACCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTG  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCGACGGCAAGGTGACCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTG  420

Qy        421 CACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCC  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCC  480

Qy        481 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGAC  540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGAC  540

Qy        541 CTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGAC  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGAC  600

Qy        601 ACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCAGGCCGGCCCGTG  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCAGGCCGGCCCCGTG  660

Qy        661 GCCCCCGGGCCGCCGAGATGCCGCGACCCCCGCGGGGCCCGGGCCGCCGCCGCACCCTG  720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCCCCGGGCCGCCGAGATGCCGCGACCCCCGCGGGGCCCGGGCCGCCGCCGCACCCTG  720

Qy        721 CAGGACAGATCGCCTCGGATGACCAGCAACCCCGCCCGTGCCGTGGGGGCGGACATTACAAG  780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAGGAGCAGATCGCCTCGGATGACCAGCAACCCCGCCCGTGCCGTGGGGCGGACATTACAAG  780

Qy        781 CGGTGGATCATCCTTGGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTG  840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGGTGGATCATCCTTGGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTG  840

Qy        841 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC  900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC  900

Qy        901 CTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTGGTG  960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTGGTG  960

Qy        961 CAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCCCTGGGCCCCGGGGCCACCCTG  1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCCCTGGGCCCCGGGGCCACCCTG  1020

Qy       1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTG  1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTG  1080

Qy       1081 GCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAGAAGAGCAACTTCAAG  1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAGAAGAGCAACTTCAAG  1140

Qy       1141 GGCCCCCGCGGCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATGGCCAAGAAC  1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGCCCCCGCGGCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATGGCCAAGAAC  1200

Qy       1201 TGCCGCGCCCCCAAGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG  1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCCGCGCCCCCAAGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG  1260

Qy       1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC  1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC  1320

Qy       1321 CCCGGCAACTTCCTGCAGAACCGCAGCGAGCCCACCGCGCCCCCGCAGCCCCAAGGAC  1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CCCGGCAACTTCCTGCAGAACCGCAGCGAGCCCACCGCGCCCCCGCAGCCCCAAGGAC  1380

Qy       1381 CCCGGCAGGCTTCGGCGCCCCGTCGAGGAGGAGACACCCCCAAGGAC  1440
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCCGGCGAGAGCTTCGGCGCCCCGTCGAGGAGGAGACACCCCCAAGGAC  1440

Qy       1441 CGCGGAGCCCTACCCGCGGAGCCCCTGACCGGCCCTGTTCGGCGAGCGGCCCCCTG  1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGCGGAGCCCTACCCGCGGAGCCCCTGACCGGCCCTGTTCGGCGAGCGGCCCCCTG  1500

Qy       1501 AGCCAGTAA  1509
              |||||||||
Db       1501 AGCCAGTAA  1509
```

RESULT 4
ABL39958

```
ID   ABL39958 standard; DNA; 1509 BP.
XX
AC   ABL39958;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Synthetic Gag polynucleotide sequence SEQ ID NO:21.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
FH
PN   WO020204493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J,  Barnett SW,  Engelbrecht S,  Van Rensburg BJ;
XX
DR   WPI; 2002-154920/20.
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Example 1; Fig 6; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptide. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (i). (i) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 1509 BP; 321 A; 560 C; 470 G; 158 T; 0 U; 0 Other;
```

```
Query Match             98.8%;  Score 1491.4;  DB 6;  Length 1509;
Best Local Similarity   99.3%;  Pred. No. 5.5e-175;
Matches 1498;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;
```

```
Qy      1 ATGGGCGCCGCGCCAGCATCCTGCGCGGCGGAGAAGTGGGAGAAGATCCGC 60
          ||||||||| ||||||||||||||| ||||||||||| |||||||||||||
Db      1 ATGGGCGCCCCGGCCCAGCATCCTGCGGCGGAGAAGCTGGGAGAAGATCCGC 60

Qy     61 CTGCGGCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCGGCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAG 120

Qy    121 CTGGAGGGCCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCGGAGGGCTGCC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGGAGGGCCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCGGAGGGCTGCC 180

Qy    181 ATGAAGCAGCTGCAGGCCCCGCCCCCTCGCAGACCCCCTGCAGCCTGTTACAAC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATGAAGCAGCTGCAGGCCCCGCCCCCTCGCAGACCCCCTGCAGCCTGTTACAAC 240

Qy    241 ACGGTGGCACCCTGTACTGCGGTCCGGACGCACCGACCCAAGGAGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACGGTGGCACCCTGTACTGCGGTCCGGACGCACCGACCCAAGGAGGCC 300
```

```
Db   1321 CCCGGCAACTTCCTGCAGAACCGACCGAGCCCGGCGCCCCACCGGCGCCCC 1380
Qy   1381 CCCGGCGGAGAGCTTCGGGCTTGGAGGAGGAGACCACCACCCCCCCCAAGGAC 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CCCGGCGGAGAGCTTCGGGCTTGGAGGAGGAGACCACCACCCCCCCCAAGGAC 1440
Qy   1441 CGCGGAGCCCTACCGCGCGAGCCCCTGAGCCGCGCCGCCGCCCCCCGCCCCTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CGCGGAGCCCTGACGCGCTGGTTCGGCAGCGGGCCCCCCTG 1500
Qy   1501 AGCCAGTAA 1509
          |||||||||
Db   1501 AGCCAGTAA 1509
```

  RESULT 5

  AAAS1626

ID   AAAS1626 standard; DNA; 1509 BP.
XX
AC   AAAS1626;
XX
DT   31-OCT-2000  (first entry)
XX
DE   HIV codon-optimized synthetic Gag polynucleotide.
XX
KW   Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW   DNA immunization; packaging cell line; antigen presentation; SS.
XX
OS   Human immunodeficiency virus; type C strain AP110967.
OS   Synthetic.
XX
PN   WO200039304-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;  99WO-US031273.
XX
PR   31-DEC-1998;  98US-0114495P.
PR   01-SEP-1999;  99US-0152195P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S, Zur Megede J;
XX
DR   WPI; 2000-452401/39.
XX
PT   Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env
PT   polypeptide and the polypeptide useful for immunizing a mammal especially
PT   human against HIV.
XX
PS   Disclosure; Page 104; 113pp; English.
XX
CC   Expression cassettes comprising a polynucleotide encoding antigenic type
CC   C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful
CC   in DNA immunization, generation of packaging cell lines and production of
CC   Gag- and/or Env-containing proteins. Synthetic Env and Gag expression
CC   cassettes exhibit increased potency for induction of cytotoxic T-
CC   lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-
CC   assemble into non-infectious virus-like particles which are used as a
CC   matrix for the proper presentation of an antigen entrapped or associated
CC   to the immune system of the host
XX
SQ   Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 U; 0 Other;

Query Match             98.7%;  Score 1489.8; DB 3;  Length 1509;
Best Local Similarity   99.2%;  Pred. No. 8.6e-175;
Matches 1497;  Conservative   0;  Mismatches   12;  Indels   0;  Gaps   0;

```
Qy      1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGAGAAGCTGGACAAGCTGGACAAGGATCGGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGAGAAGCTGGACAAGCTGGACAAGGAC 60
Qy     61 CTGCGCCCCGGCGGCAAGCTGGACAAGTGGGAGAAGATCCGC 120
          ||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAAGAAGCTGGTGTGGGCGGCCAGGAG 120
```

```
Db  1141 GGCCCCCGGCCCCGCCAACGTGAAGTGTCTTCAAGTCGGGCAAGGAGGAGGAGGGGCCCCACATCGCCCAAGAAC 1200
Qy  1201 TGCCCGCGGCCCCCGGCAAGAAGAGGGCTGCTGGGAGTGCGGGCCAGCAGATGAAG 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 TGCCCGCGGCCCCCGGCAAGAAGAGGGCTGCTGGGAGTGCGGGCCAGCAGATGAAG 1260
Qy  1261 GACTGCACCGAGGCGCCAGGCCCAACTTCCTTCCTGGGAGTCGGGCCAGCCACAAGGGGCGC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GACTGCACCGAGGCGCCAGGCCCAACTTCCTTCCTGGGAGTCGGGCCAGCCACAAGGGGCGC 1320
Qy  1321 CCCGGCAACTTCCTGCAGAACCGCAGGGAGCCGGAGCCGGGCCCGTGCCCACCGCCCCC 1380
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CCCGGCAACTTCCTGCAGAACCGCAGGGAGCCGGAGCCGGGCCCGTGCCCACCGCCCCC 1380
Qy  1381 CCCGGCGAGAGCTTCCGCCTTCGAGGAGACCACCCCCCCCAAGCACAGGAGGAGCCCAAGGAC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 CCCGGCGAGAGCTTCCGCCTTCGAGGAGACCACCCCCCCCAAGCACAGGAGGAGCCCAAGGAC 1440
Qy  1441 CGGCGAGCCCTACCGGCGACCCGCCCGCCCGCGCGAGCGCGAGCGCCTTGTTCCGGCAGCGGGGCCCCCCTG 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 CGGCGAGCCCTACCGGCGACCCGCCCGCCCGCGCGAGCGCGAGCGCCTTGTTCCGGCAGCGGGGCCCCCCTG 1500
Qy  1501 AGCCAGTAA 1509
              |||||||||
Db  1501 AGCCAGTAA 1509
```

```
RESULT 6
AAL44553
ID   AAL44553 standard; DNA; 1509 BP.
XX
AC   AAL44553;
XX
DT   29-AUG-2003 (revised)
DT   08-NOV-2002 (first entry)
XX
DE   Human immunodeficiency virus 1.
XX
KW   HIV-1 p55gag polypeptide coding sequence 4.
XX
OS   Human immunodeficiency virus 1.
XX
FH   Key          Location/Qualifiers
FT   variation    replace(282, C)
FT                /tag= a
FT   variation    replace(331, T)
FT                /tag= b
FT   variation    replace(332, C)
FT                /tag= c
FT   variation    replace(549, G)
FT                /tag= d
FT   variation    replace(553, T)
FT                /tag= e
FT   variation    replace(783, G)
FT                /tag= f
FT   variation    replace(816, G)
FT                /tag= g
FT   variation    replace(999, T)
FT                /tag= h
FT   variation    replace(1002, C)
FT                /tag= i
FT   variation    replace(1089, G)
FT                /tag= j
FT   variation    replace(1149, G)
FT                /tag= k
FT   variation    replace(1158, C)
FT                /tag= 1
```

```
XX
PD   04-APR-2002.
XX
PD   28-SEP-2001; 2001WO-US030540.
XX
PF   28-SEP-2000; 2000US-0236105P.
PR   30-AUG-2001; 2001US-0315905P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   O'hagan D, Otten G, Donnelly JU, Polo JM, Barnett S, Singh M;
PI   Ulmer J, Dubensky TW;
XX
DR   WPI; 2002-519084/55.
XX
PT   HIV; dsi; vaccine; gene; immune response; microparticle;
PT   adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
PT   polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
PT   submicron emulsion; viral infection; bacterial infection;
PT   parasitic infection; HIV-1 p55gag polypeptide.
XX
PS   Claim 72; Fig 6; 100pp; English.
XX
CC   The invention relates to a method of raising an immune response in a host
CC   animal. The method of the invention comprises administering a
CC   microparticle that has an adsorbent surface to which a first biologically
CC   active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC   microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC   acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC   a polycyanoacrylate, a detergent, and submicron emulsion. The method/
CC   microparticle of the invention is useful for immunising a host animal
CC   against viral, bacterial or parasitic infections. The present DNA
CC   sequence encodes a HIV-1 p55gag polypeptide. [Updated on 29-AUG-2003 to
CC   standardise OS field]
XX
SQ   Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 U; 0 Other;
```

```
Query Match         98.7%;  Score 1489.8;  DB 6;  Length 1509;
Best Local Similarity   99.2%;  Pred. No. 8.6e-175;
Matches 1497;  Conservative 0;  Mismatches 12;  Indels 0;  Gaps 0;
```

```
Qy    1 ATGGGGCGCCCGGCCGCCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGGGGCGCCCGGCCGCCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 60
Qy   61 CTGCGCCCCGGCGGGCAAGAGAAGCACTACAATCTGGTGTGGGCCAGCCGCGCGGAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGCGCCCCGGCGGGCAAGAGAAGCACTACAATCTGGTGTGGGCCAGCCGCGCGGAG 120
Qy  121 CTGGGAGGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACACCGGCCGGAAGCTGGCAGAGATC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CTGGGAGGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACACCGGCCGGAAGCTGGCAGAGATC 180
Qy  181 ATGAAAGCAGCTGCAGCCGCAGCCGCCCATGTGCAGACCACCGAGCCTGCGCAGCCTGTACAAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ATGAAAGCAGCTGCAGCCGCAGCCGCCCATGTGCAGACCACCGAGCCTGCGCAGCCTGTACAAC 240
Qy  241 ACCGTGGCCACCCTGTACTGCGTGCACAGCAGGGTCCGGACAGCCAAGGAGGAGCC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ACCGTGGCCACCCTGTACTGCGTGCACAGCAGGGTCCGGACAGCCAAGGAGGAGCC 300
Qy  301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAG 360
Qy  361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTG 420
Qy  421 CACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 CACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCC 480
Qy  481 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCGCCCCTGAGCGAGGGCGCCACCCCCCAGGAC 540
```

```
PN   WO200226209-A2.
```

RESULT 7

```
Db   481 TTCAGCACCCGAGGTGATCCCATGTTCACCGCCATCGCCCTGAGCGAGGCGGCCACCCCCAGGAC
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541 CTGAACACGAGGAGGCTTGAACACCGGTGGGGGGGGGGGGGCCCCCATCAGATGCTGAGGAC 600
Db   541 CTGAACACCATGCTGAACACCGGTGGCGGCGGCGGCCCCCATGCAGATGCTGAAGGAC 600
Qy   601 ACCATCAACGAGGAGGCCGGCCGCGAGGTGGGACCGCCGAGGTGCAGGGCGGCCCCGTG 660
Db   601 ACCATCAACGAGGAGGAGGCCGCCGGCGGGTGGGACCCTCGCACCCGGTGCAGGGCGGCCCCGTG 660
Qy   661 GCCCCCCGGCCAGTGCCGGACCCCCCGGACCAGACACCGGCGGCCGCCACCCCGTG 720
Db   661 GCCCCCCGGCCAGTGCCGGACCCCCCGGACCAGCGGCCGCCACCCCCTG 720
Qy   721 CAGGAGCGAGATCGGCCTGGGATGACCACCGGCAACCCCCCGCGGCGGCGACCATCTACAAG 780
Db   721 CAGGAGCAGATCGGCCTGGGATGACCACCGAACCCCCCGGGCGGCGACCATCTACAAG 780
Qy   781 CGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTG 840
Db   781 CGCTGGANTCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTG 840
Qy   841 GACATCAGGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAGACC 900
Db   841 GACATCCGGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAGACC 900
Qy   901 CTGCGCGCGGAGCAGGCCACCCCAGAGGTGAAGAACTGGATGACCGAGACCCTGCTGGTG 960
Db   901 CAGAACCCCAACCCCGACTGCAAGACCATCCTGCGGGCGCTTCTGGCGGACCCCCACCCTG 1020
Qy   961 CAGAACCCCAACCCCGACTGCAAGANCCATCCTGCGGGGCCTCTGGCGGACCCCCACCCTG 1020
Db   1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTG 1080
Qy   1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTG 1080
Db   1081 GCCGGAGCCATGAGCCAGGCCAACAGCGGTGAACATCATGATGCAGAGAGCAACTTCAAG 1140
Qy   1081 GCCGGAGGCCATGAGCCAGGCCAACAGCGGTGAACATCATGATGCAGAGAGCAACTTCAAG 1140
Db   1141 GGCCCCCGGCGCACCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAAC 1200
Qy   1141 GGCCCCCGGCGCACCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAAC 1200
Db   1201 TGCCGCGCCCCCCGGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
Qy   1201 TGCCGCGCCCCCCGGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
Db   1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCCACCACCAAGGGCCGC 1320
Qy   1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCCACCACAAGGGCCGC 1320
Db   1321 CCCGGCAACTTCCTGCAGAACCGCCCGGAGCCGAGAAGTTGGACAAGTGGAGAGAGATCCGC 1380
Qy   1321 CCCGGCAACTTCCTGCAGAACCGCCCGGAGCCGAGAAGTTGGACAAGTGGAGAGCGCATCCGC 1380
Db   1381 CCCGGCCAGAGCTTCGGCTTCGGAGGAGGACCAGCAGGAGCCAGAC 1440
Qy   1381 CCCGGCCAGAGCTTCGGCTTCGGAGGAGGACCAGCAGGAGCCAGAC 1440
Db   1441 CGCGGAGCCCTACCGCGAGCCCTGACCGCCCTGTTCGGGCAGCCCCTG 1500
Qy   1441 CGCGGAGCCCTACCGCGAGCCCTGACCGCCCTGTTCGGGCAGCCCCTG 1500
Db   1501 AGCCAGTAA 1509
Qy   1501 AGCCAGTAA 1509
```

ABL39972

Query Match             86.0%;  Score 1297.6;  DB 6;  Length 1494;
Best Local Similarity    92.7%;  Pred. No. 3.3e-151;
Matches 1402; Conservative   0; Mismatches  89; Indels  21; Gaps  3;

```
Qy     1 ATGGGCGCGCCGGCGCCGCCCAGCATCCTGCGGGCGGGAGAAGTTGGACAAGTGGAGAGATCCGC 60
         |||||    ||||||||||||||||||||||||||||||||||||||||||||     ||
Db     1 ATGGGCGCGCCCGGCGCCGCAGCATCCTGAGCGCGGCCAAGCTCGGACAAGTGGGAGCGCCGCAT 60
Qy    61 CTGGCGGGTCCCGGGCGGCAAGAAGCACTACATGCTGAAGGCACCTGGTGTGCGGCCAGCGCGAG 120
         |||||| ||||||||||||||||||   |||||||||||||||||||
Db    61 CTGGCGGCCCCGGGCGGCAAGAAGCACTACATGCTGAAGGCACCTGGTGTGCGGCCAGCGCGAG 120
Qy   121 CTGGAGAGGGGCTTCGCCCTGAACCCGGCCTGCTGAGACCGCGACCGGAGCGAGCCTGTACAAC 180
         |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGGAGACGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
Qy   181 ATGAAGCAGCTGCAGCCCTCCCTGCAGACCGGCAGCGAGGAGCTGCGCAGCCTGTACAAC 240
         |||| ||||||||||||||||||||||||||||||||||||||||||  |||| ||||
Db   181 ATCAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTTCAAC 240
Qy   241 ACCGTGGCCACCCTGTACTGCGTGCACGCCCGCATCGAGGTCCGGGACACCACCAAGGAGGCC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 8
ABL40020
ID ABL40020 standard; DNA; 1491 BP.
XX
AC ABL40020;
XX
DT 15-MAY-2002 (first entry)
XX
DE Synthetic Gag polynucleotide sequence SEQ ID NO:99.
XX
KW Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX
OS Human immunodeficiency virus; type C.
OS Synthetic.
XX
FH
RN WO20204493-A2.
XX
PD 17-JAN-2002.
XX
PF 05-JUL-2001; 2001WO-US021241.
XX
PR 05-JUL-2000; 2000US-00610313.
XX
PI zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
DR WPI; 2002-154920/20.
XX
PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT applications including DNA immunization or generation of packaging cell
PT lines, particularly in gene therapy.
XX
PS Claim 8; Fig 70; 233pp; English.
XX
CC The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC type C polypeptides. The expression cassettes comprise any of the HIV
CC type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC (I). (I) have immunostimulant activity and can be used in gene therapy.
CC The HIV type C polynucleotides are useful in applications including DNA
CC immunisation, generation of packaging cell lines, and production of HIV
CC Type C proteins. The polynucleotides are particularly useful in gene
CC therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC ABB06204 to ABB06215 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 1491 BP; 319 A; 557 C; 457 G; 158 T; 0 U; 0 Other;

Query Match 85.8%; Score 1294.8; DB 6; Length 1491;
Best Local Similarity 92.7%; Pred. No. 7.3e-151;
Matches 1401; Conservative 0; Mismatches 87; Indels 24; Gaps 3;

Qy 1 ATGGGCGCCCGGCGCCAGCATCCTGCGCGGGACGCTGGACAAGTGGGAGAAGATCCGC 60
Db 1 ATGGGGCCCCGGCGCCAGCATCCTGCGCGGGCAAGCTGGACAAGTGGGAGAGAAGATCCGC 60

[Sequence alignment blocks (Qy / Db pairs, residues 241–1509) appear in the left portion of the page as dense nucleotide alignment data.]

```
Qy    61 CTGCGCCCCGGCGGGGGGCAAGAGAGACTACATGCGTAGTGAAAGCACCGGTGTGTGGGCCAGCGCGGAG 120
         ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    61 CTGCGCCCCGGCGGGGGGCAAGAGAGACTACATGCGTAGTGAAAGCACCGGTGTGTGGGCCAGCGCGGAG 120

Qy   121 CTGGGAGGGCCTTGGCCCTGCTCTGAGAACCCCGGCGGCTGCAAGCAGATC                     180
         ||||||||||||||||||||      ||| ||||||||||||||
Db   121 CTGGGAGGGCCTTGGCCCTGCGTGAACCCCGGCGGCTGCAAGCAGATC                        180

Qy   181 ATGAAGCAGCTGCAGCCCGCGCTGCAGACGGGCACCGAGGAGCTGCGCAGCCTGTACAAC            240
         ||||||||||||||||||||||||||||||||||||||
Db   181 ATCAAGCAGCTGCAGCCCGCGCTGCAGACGGGCACCGAGGAGCTGCGCAGCCTGTTCAAC            240

Qy   241 ACCGTGGCCACCCTGTACTGCGTGCACGCCGGGATCGAGGTCCGGGCACACCAGAAGAGGCC          300
         |||||||||||||||||||||||||||||||||
Db   241 ACCGTGGCCACCCTGTACTGCGTGCACGCCGAGATCGAGGTGCGGGACACCAAGGAGGCC            300

Qy   301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCGAGGCC            360
         ||||||||||||||||||||||||||||||
Db   301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGACCCAGCAGGCCGAGGCC            360

Qy   361 GCCGGAACGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTG        420
         ||||||||||||||||
Db   361 GCCGGACAAGAGAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTG        420

Qy   421 CACCAGGCCATCAGCCCGAGGACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCC          480
         ||||||||||||||||
Db   421 CACCAGGCCATCAGCCCGAGGACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCC          480

Qy   481 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGAC          540
         ||||||
Db   481 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGAC          540

Qy   541 CTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGAC          600
         |||||
Db   541 CTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGAC          600

Qy   601 ACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCACGCCGGCCCGGTG          660
         ||||||||||||||||
Db   601 ACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCACGCCGGCCCGGTG          660

Qy   661 GCCCCCGGCCAGATGCGGGACCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACCCTG          720
         |||||||
Db   661 GCCCCCGGCCAGATGCGGGACCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACCCTG          720

Qy   721 CAGGAGCAGATCGCCTTGGATGACCAGCAAACCCCGTCCCGTGGGCGCAACATCTACAAG          780
         |||||
Db   721 CAGGAGCAGATCGCCTTGGATGACCAGCAAACCCCGTCCCGTGGGCGCAACATCTACAAG          780

Qy   781 CGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTG          840
         |||||||||||||||||
Db   781 CGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTG          840

Qy   841 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAGACC          900
         ||||||
Db   841 GACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAGACC          900

Qy   901 CTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTGGTG          960
         |||||||
Db   901 CTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTGGTG          960

Qy   961 CAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGCGCCACCCTG         1020
         ||||||||
Db   961 CAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGCGCCACCCTG         1020

Qy  1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTG         1080
         |||||||||||
Db  1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTG         1080

Qy  1081 GCCGGAGGCGGCATGAGCCAGAGCCAACAACACCAGCGGTGACATCATGATGTCGCAGAACTTCAAG 1140
         ||||||||||||| |
Db  1081 GCCGGAGGCGGCATGAGCCAGAGCCAACAACACCAGCGGTGACATCATGATGTCGCAGAACTTCAAG 1140

Qy  1141 GGCCCCCGCCGCGCAAGTCAAAGTGCTTCAACTGCGGCAAGGAGGGGCCCACATTCGCCAAGAAC     1200
         |||| |||   |||
Db  1141 GGCCCCCGCCGCGCGCGCCGTGAAGTGCTTCAACTGCGGCAAGGAGGGGCCCACATCGCCAAC       1200

Qy  1201 TGCCGGCGCCCCCGGGACCCCCAAGCGGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCACCAGATGAAG 1260
         |||||||||
Db  1201 TGCCGGCGCCCCCGGGACCCCCAAGCGGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCACCAGATGAAG 1260

Qy  1261 GACTGCACCGAGCGCCAGGCGACCTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC         1320
         |||||||
Db  1261 GACTGCACCGAGCGCCAGGCGACCTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC         1320

Qy  1321 CCCGGCAACTTCCTGCAGAACCGCCCGGAGCCCACCGCCCCGCCCGAGGAGAGCTTCCGC         1377
         |||||||||||
Db  1321 CCCGGCAACTTCCTGCAGAACCGCCCGGAGCCCACCGCCCCGCCCGAGGAGAGCTTCCGC         1380

Qy  1378 CCCCGGCCGGAGAGCTTCAAGTTCGAGGAGA------CCCGCAAGGAGAGGGGCCCAAG     1437
         |||||||
Db  1381 CCCCGGCCGGAGAGCTTCAAGTTCGAGGAGACCCGCCCCCGCAAGGAGAGGGGCCCAAG     1431

Qy  1438 GACCCGCGGAGGCCCTACCGGCCGCCCGTGACCCGCCGCGCAGCCTTCGGCGGGCGGCCC      1497
         ||||||||||||
Db  1432 G-------ACCGGCGGAGCCCCTGACCAGCGCCTTGAAGAGGCCTTCGGCGGCCGGCCC      1479

Qy  1498 CTGAGCCACTAA 1509
         |||||||||
Db  1480 CTGAGCCAGTAA 1491
```

PS Claim 2; Page 92-93; 113pp; English.
XX
CC Expression cassettes comprising a polynucleotide encoding antigenic type
CC C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful
CC in DNA immunization, generation of packaging cell lines and production of
CC Gag- and/or Env-containing proteins. Synthetic Env and Gag expression
CC cassettes exhibit increased potency for induction of cytotoxic T-
CC lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-
CC assemble into non-infectious virus-like particles which are used as a
CC matrix for the proper presentation of an antigen entrapped or associated
CC to the immune system of the host
XX
SQ Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Query Match 84.6%; Score 1276.8; DB 3; Length 1479;
Best Local Similarity 92.5%; Conservative 0; Mismatches 77; Indels 36; Gaps 4;
Matches 1399;

```
Qy      1  ATGGGGCGCCGGCCGCCAGCCATCCTGGGGCGGCCGGAGAAGCTTGGACAAGTGGGAGATCCGC   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGGGGCGCCGGCCGCCAGCCATCCTGGGGCGGCCGGAGAAGCTTGGACAAGTGGGAGATCCGC   60

Qy     61  CTGGCCCCGGGCGGAGAAGCACTACATGCTGAAGCAACCTGGTGTGTGGCCAGGCCGCATCCGAG  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CTGCGCCCCGGGCGGAGAAGTGCTACATGGTTGGGTGTGGGCCAGCCGCATCCCGGAG  120

Qy    121  CTGGAGGGCGCTTCGCGCCCTGAACCCCCGGCCTTGCTGGGACACCCGGCCTGCAAGCCAGATC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  CTGGGAGAAGTTCGGCCCTGAACCCGGCCTTGCTGGGACACCCGGCCTGCAAGCCAGATC  180

Qy    181  ATGAAGACGCAGCTACAGCCCCCTCGAGACGGAGCTCGCACCCTGTCAACC  240
           |||||||||||||||||||||||||||||||||||||||||||||
Db    181  ATCCGCACAGCTGCACCCCCGGCCCTGCAGACCGGAGCTGAAGAGCCTGTTCAAC  240

Qy    241  ACCGTGGCCACCCTGTACTGCGTGCATCAGAGGTTCGGGACACCAAGGAGGCC  300
           |||||||||||||||||||||||||||||
Db    241  ACCGTGGCCACCCTGTACTGGTGCGTGCATCAGGAGATCGGAGGTTCGGAGGGCC  300

Qy    301  CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAAACCCCACCAGGATC  360
           |||||||||||||||||||||||
Db    301  CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCCCGAGGGC  360

Qy    361  GCCGGAC---GGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAAGGGCCCAGATG  417
           ||||||||       |||||||||||||||||||||||||||
Db    361  GCCCAGACAAGGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGGCCCAGATG  417

Qy    421  GTGCCACCAGGGCCATCAGCCCCGCCCTGAACCCTGGGTGTGAAGGTGATGCAGGAGAG  477
           |||||||||||||||||||||||||||||||||||||||
Db    421  GCCCACCAGGCCAGCAGGCCATCCCCGGGACTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG  480

Qy    478  GCCTTCAGCCCCGAGGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGCGCCACCCCCCAG  537
           |||||||||||||||||||||||||||||
Db    481  GCCTTCAGCCCCGAGGTGATCCCCATGTTCACAGCCCTGAGCGAGGGCGCCACCCCCCAG  540

Qy    538  GACCTGAACACGATGTTGAACACCGTGGGCGGGCCGCCACCGGAGCAGCTGCTGAG  597
           |||||||||||||||||||||||
Db    541  GACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG  600

Qy    598  GACACCATCAACGAGGAGGCCGCCGAGTGGGACAGGGTGCACCCGGTGCACGCAGGCCGGCCCC  657
           |||||||||||||||||||||||||||
Db    601  GACAACATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCCATC  660

Qy    658  GTGGCCCCCGGCCAGATGCGCGAGCCCAGGGACCGCCCAGGACAGGCACC  717
           ||||||||||||||||||||||||||||||
Db    661  ATCGCCCCCGGCCAGATGCGCGAGCCCAGGGGCAGCGACATCGCCGGCACC  720

Qy    718  CTGCAGGAGGCAGAGTCGGCGGCTGGATGACAGCACCAGACAGACTCTAC  777
           ||||||||||||||||||||
Db    721  CTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCTGTGGGCGACATCTAC  780

Qy    778  AAGGGGGTGGAATCATCCTGGGCCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC  837
           ||||||||||||||||||||||
Db    781  AAGCGGGTGGGATCATCCTGGGCCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC  840
```

```
Qy    838  CTGGACATCCGGCCAGGCGCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCGTTCTTCAAG  897
           |||||||||        |||||||||||||||||||||||||||||||||||||||||||
Db    841  CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG  900

Qy    898  ACCCTGCGCGCCGAGCAGGCCAGCCAGGAGGTGAAGAACTGGATGACGGAGACCCTGCTG  957
           ||||||||||||||||||||||||||||||||||||||
Db    901  ACCCTGCGCGCCGAGCAGGCCAGCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG  960

Qy    958  GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGGGCCCCGCCGCCACC  1017
           |||||||||||||||||||||||||||||||||||||
Db    961  GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGGGCCCCGCCGCCACC  1020

Qy   1018  CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG  1077
           |||||||||||||||||||||||||||||||||||||
Db   1021  CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG  1080

Qy   1078  CTGGCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAGAAGAGCAACTTC  1137
           |||||||||||||||||||||||        |||||||||||||||||
Db   1081  CTGGCCGAGGCCATGAGCCAGGCCACCAGCAGCGTGAACATCATGATGCAGAAGAGCAACTTC  1137

Qy   1138  AAGGGCCCCCGGCCGCCAACGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAG  1197
           |||||||||||||||||||||||||||||||||||||
Db   1138  AAGGGCCCCCGGCCGCCAACGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGC  1197

Qy   1198  AACTGCCGCGCCCCCCGGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG  1257
           |||||||||||||||||||||||||||||||||||||
Db   1198  AACTGCCGCGCCCCCCGGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG  1257

Qy   1258  AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC  1317
           |||||||||||||||||||||||||||||||||||||
Db   1258  AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC  1317

Qy   1318  CGCCCCGGCAACTTCCTGCAGAACCCCGAGCCCACCGCTCCCCCAGCCCCGAGCCCACCGCC  1377
           ||||||||||||||||||||||||        |||||||
Db   1318  CGCCCCGGCAACTTCCTGCAGAACAGGCCCGAGCCCACCGCCCCCCCCGAGGCCAGCCAGTTCCGC  1377

Qy   1378  CCCCCCGCCGGAGAGCTTCGGCGTTCGAGGAGACCACCCCCCCGCCCCCGGCCCAGCCCCCAAG  1437
           |||||||||||||||||||||||||||||||||||||
Db   1378  CCCCCCCGGCCGGAGCTTCGGCTTCGAGGAGACCACCCCCCCGCCCCCGGCCCAGCCCCCAAG  1419

Qy   1438  GACCGCGAGCCCTACCCCGAGCCCGCCTGACCTGCTGGGCCAGCCGCCCCGGGCCCC  1497
           ||||||||||||||        |||||||||||||||||
Db   1420  G-------ACCCGGAGACCCCAGCCTGAAGAGCCTGTTCGGGCCAACGACCCC  1467

Qy   1498  CTGAGCCAGTAA  1509
           ||||||||||||
Db   1468  CTGAGCCAGTAA  1479
```

PF   28-SEP-2001; 2001WO-US030540.
XX
PR   28-SEP-2000; 2000US-0236105P.
PR   30-AUG-2001; 2001US-0315905P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;
PI   Ulmer J, Dubensky TW;
XX
DR   WPI; 2002-S19084/55.
XX
PS   Claim 72; Pig 1; 100pp; English.
XX
PT   The invention relates to a method of raising an immune response in a host
PT   animal. The method of the invention comprises administering a
CC   microparticle that has an adsorbent surface to which a first biologically
CC   active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC   microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC   acid, a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC   a polyanhydride, a polycyanoacrylate, and submicron emulsion. The method
CC   of the invention is useful for immunising a host animal.
CC   microparticle of the invention is useful for immunising a host animal
CC   against viral, bacterial or parasitic infections. The present DNA
CC   sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to
CC   standardise OS field)
XX
SQ   Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Query Match   84.6%;   Score 1276.8;   DB 6;   Length 1479;
Best Local Similarity   92.5%;
Matches 1399; Conservative   0;   Pred. No. 1.2e-148;
  Mismatches 77;   Indels 36;   Gaps 4;

A microparticle to which a biologically active macromolecule is adsorbed,
for use as a vaccine composition to treat viral, bacterial or parasitic
infections, comprises a polymer microparticle, a detergent and a.
submicron emulsion.

RESULT 11
ABL39953
ID   ABL39953 standard; DNA; 1479 BP.
XX

Qy     1 ATGGGCGCCGCGGCCCAGCCATCCTCGCGGGAGAGAGTGGACAAGTGGGAGAAGATCCGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGGGCGCCGCGGCCCAGCCATCCTCGCGGGAGAGAGTGGACAAGTGGGAGAAGATCCGC 60

Qy    61 CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CTGCGCCCCGGCGGCAAGAAGTACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120

Qy   121 CTGGAGAGGGCCTTCGCCCTGCCTGCTTGGGAGACCCGCGACGCTGCCAAGCACAATC 180
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGGAGAAGTTCGCCCTCAAACCCGGGCCTCCTGGAGACCCAGCGACGCTGCCAAGCACAATC 180

Qy   181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACGGGCCACCGGCGAGCTGCGCAGCTGCTGCTACAAC 240
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ATCCGCCAGCTGCACCCCGCCCTGAGCACCGGCGAGGCTGTTCAAC 240

Qy   241 ACCGTGGCCCACCCTGTACTGCGGTGCACCGGGCCATCGAGGTCCGGGCC 300
     |||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ACCGTGGCCCACCCTGTACTGCGGTGCACCGGGCCATCGAGGTCCGGGCC 300

Qy   301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGCCCAAGGAG 360
     |||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGCCCAAGGAGGCC 360

Qy   361 GCCGAC---GGCAAGGTGAGCCAGAACTACCCCATCGTTCAGAACATTCCAGGGCCCAGATG 417
     |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCGACAAGCAGTGAGCCAGAACTACCCCATCGTTCAGAACATTCCAGGGCCCAGATG 420

Qy   418 GTGCACCAGGCCCATCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCCCAG 477
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTGCACCAGGCCCATCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCCCACCCCCCAG 480

Qy   478 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCCCAG 537
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCCCACCCCCCAG 540

Qy   538 GACCTGAACACGATGTTGAACACGTGGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 597
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GACCTGAACACGATGTTGAACACGTGGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600

Qy   598 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCGGTGCAGCCCGCCC 657
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCGGTGCAGCCCGCCC 660

Qy   658 GTGGGCCCCCGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 717
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ATCCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 720

Qy   718 CTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTGGGCGACATCTAC 777
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTGGGCGACATCTAC 780

Qy   778 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 837
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840

Qy   838 CTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 897
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900

Qy   898 ACCCTGCGCGCCGAGCAGGCCGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTG 957
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 ACCCTGCGCGCCGAGCAGGCCGCCACCCAGGAGGTGAAGAACTGGATGACCGAGACCCTGCTG 960

Qy   958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGGGCCCCGCCGCCACC 1017
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCCCTGGGCCCCGCCGCCAGC 1020

Qy   1018 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAGAAGGCCCGCCTG 1077
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAGAAGGCCCGCCTG 1080

Qy   1078 CTGCGCCGGGAGGCGGATGAGCCAGCAGCAGCCCAACACGGTGAACATCATGATGCAGAGAGAGCAACTTC 1137
     ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   1081 CTGGGCGCCGGATGAGCCAGCAGCAGCCCAACAC---AGCGGTGATGATGCAGAGAGAGCAACTTC 1137

Qy   1138 AAGGGCCCCCGCCGGGGCAAACGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAG 1197
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1138 AAGGGCCCCCGCCGGGGCAAACGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAG 1197

Qy   1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257

Qy   1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1317
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1317

Qy   1318 CGCCCCGGCAACTTCCTCCAGAGCCG-------CCCCGAGCCCACCGCC 1359
     |||||||||||||||||||||||||       ||||||||||||||||||
Db   1318 CGCCCCGGCAACTTCCTGCAGAGCAGGCCCGAGCCCACCGCC 1377

Qy   1360 CCCCCGCCCGAGGAGTTCCGCTTCGGAGAAGACCAGCAGAGGACGCCAAG 1419
     ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1378 CCCCCGCCCGAGGAGTTCCGCTTCGGAGAAGAGACCAGCAGAGGACGCCAAG 1437

Qy   1420 G------------ACCGCGAGACCCTGACCAGCCTGCGCAGCCTGTTCGGCAACGACCCC 1467
                  |||||||||||||||||||||||||||||||||||||||
Db   1438 GACCGCGAGGACCCTGACCAGCCTGCGCAGCCTGTTCGGCAACGACCCC 1497

Qy   1468 CTGAGCCAGTAA 1479
     ||||||||||||
Db   1498 CTGAGCCAGTAA 1509

```
AC   ABL39953;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Synthetic Gag polynucleotide sequence SEQ ID NO:3.
XX
DB
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
OS
XX
PN   WO20040493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SM, Engelbrecht S, Van Rensburg BJ;
XX
DR   WPI; 2002-154920/20.
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Example 1; Fig 1; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Query Match             84.6%;  Score 1276.8;  DB 6;  Length 1479;
Best Local Similarity   92.5%;  Pred. No.1.2e-148;
Matches 1399; Conservative    0; Mismatches   77;  Indels   36;  Gaps    4;

Qy          1 ATGGGGCGCCCCGGCCAGCATCCTGCGCGCGCGCCAGCAAGTGGGAGAAGATCCGC   60
              |||||||||||||||||||||||||||||||||||||||||||| | |||||||||
Db          1 ATGGGGCGCCCCGGCCAGCATCCTGCGCGCGCGCCAGCAAGCTGGGAGCGCATCGC   60

Qy         61 CTGCGCCCCCGCGGCAAGAAGCACTACATGCTGAAGCATCTGGTGTGGGCCAGCCGCGAG  120
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db         61 CTGCGCCCCCGCGGCAAGAAGCACTACATGCTGAAGCATCTGGTGTGGGCCAGCCGCGAG  120

Qy        121 CTGGAGAGGTTCGCCGTGAACCCCGGCCTGCTGGAGACGAGCGAGGGCTGCAAGCAGATC  180
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGAGAAGTTCGCCGTGAACCCCGGCCTGCTGGAGACCAGCGAAGCCCTGGACAGATC  180

Qy        181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCCGCCAGTTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC  240

Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTCCGGCACACCAAGGAGGCC  300
              |||||||||||||||||||| |||| |||||||||||||||||||||||||||||||||
Db        241 ACGGTGGCCACCCTGTACTGCGTGCACGAGAAGATGAAGGTCCGCGACACACCAAGGAGGCC  300

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCACCAGAGAAGACCCAGCAGCCAAGGAG  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGAAGCAGGAAGCTGCAGGCC  360

Qy        361 GCCGCAC---GGCAAGGTGGGCGCAGGAACTACCCCATCGTTGCAGCGACATCGCAGATG  417
              |||||||   |||||||||||||||||||||||||||||||||||||||||| |||||||
Db        361 GCCGACAAGGCCAGGTGAGCCAGAACTACCCCATCGTCCAGAACCTGCAGGGCCAGATG  420

Qy        418 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG  477
              |||||||||||| | || || |||||||||||||||||||||||||||||||||||||||
Db        421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG  480

Qy        478 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG  537
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG  540

Qy        598 GACACCATCAACGAGGAGGCCGCCGAGTGGGACAGGCTGCACCCCGTGCACGCCGGCCCC  657
              |||||||||||||||||||||||||||||| ||||||||||||| ||||||||||||||
Db        541 GACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG  600

Qy        658 GTGGCCCCCGGCCAGATGCGCGACCCCCGCGGCAGCGACGACATCGCCGGCACCACCAGC  717
              |||||||||||||||||||||||||||||||||||||||||| || |||||||||||||
Db        601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACAGGCTGCACCCCGTGCACGCCGGCCCC  660

Qy        718 ATCGCCCCCGGCCAGATGCGGGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC  777
              |||||||||||||||||| || ||||||||||||||||||||||||||| | ||| ||
Db        661 ATCGCCCCCGGCCAGATGCGGGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC  720

Qy        778 AAGCGGCGTGGATCATCCTGGGCCTTGAAACAAGATCGTGCGGATGTACAGCCCCGTGAG  837
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTGGGCGACATCTAC  780

Qy        838 CTGGACATCCGGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGTTCTTCAAG  897
              |||||||| || ||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC  840

Qy        898 ACCCTGCGGGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTG  957
              |||| | |||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG  900

Qy        958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCGCTGGGCCCGGGCGCCACC  1017
              |||||||| |||||||||||||||||||||| ||| | |||||||||||||||||||||
Db        901 ACCCTGCGCGCCGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGAGACCCTGCTG  960

Qy        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTGGGCCCGGGCGCCACC  1020
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTGGGCCCGGGCGCCACC  1020

Qy       1018 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG  1077
              |||||||||| ||| |||||||||||||||||||||||| |||||||||||||||||||
Db       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG  1080

Qy       1078 CTGGCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAGAAGAGCAACTTC  1137
              ||||| ||||| |||||||||||||||||   |   ||||||||||||||||||||||||
Db       1081 CTGGCCGAGGCGATGAGCCAGGCCAACCC---AGCGGTGAATGCAGAAGAGCAACTTC  1137

Qy       1138 AAGGGCCCCCGCCGGACCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAG  1197
              ||||||||||| |||||||||||||||||||||||||||||||||| ||| |||||||||
Db       1138 AAGGGCCCCCGCCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAG  1197

Qy       1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG  1257
              |||||||||||||||| |||||||||||||||||||| |||||||||||||||||||||||
Db       1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG  1257

Qy       1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCAACAAGGGC  1317
              ||||||||| |||||||||||||||| |||||| | |||||||| ||||||||||||||
Db       1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCAACAAGGGC  1317

Qy       1318 CGGCCCGGCAACTTCCTGCAGAACCGCCCGGAGCCCACCGCCCCGCCCGTGCCCACCGCC  1377
              ||||||||| |||| |||||||||||||| ||| |||||||||||||||||||||||||
Db       1318 CGGCCCGGCAACTTCCTGCAGAGCCGG------CCCGGAGCCCACCGCCCCGCCCGTGCCCACCGCC  1359

Qy       1378 CCCCCGCCGGAGAGCTTCGGGAGGGACCCACCCCCCAAGCAGGAGGCCAAG  1437
              ||||| | |||||||||||||||||||||||||||||||||||||||||||
Db       1378 CCCCCGCCGGACCCACTTCCTGGCGACAGAGGCG  1359
```

```
Db   1360 CCCCCCCCGAGGAGCTTCGGCTTCGGAGGAGGACCACCCCGGCGGGCCTCCAGAAGCGGAGGAGCAAG 1419
          ||||||||||||||||||||||||||
Qy   1438 GACCGCGGAGCCCTACCCGGAGCCCTGGGCAGCCTGTTGTTCGGCACGCGGCCCC 1497
          |||||||||||||           ||||||||||||||||||
Db   1420 G--------------ACCCGGAGCCCTGAACGCCTGTTCGGCAACGACCCC 1467

Qy   1498 CTGGAGCCAGTAA 1509
          |||||||||||||
Db   1468 CTGGAGCCAGTAA 1479
```

RESULT 12
ADC13239

```
ID   ADC13239 standard; DNA; 3162 BP.
XX
AC   ADC13239;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagRTmut_C SEQ ID NO 18.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO200304620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J,  Barnett SW,  Lian Y,  Engelbrecht S,  Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Pig 15; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3162 BP; 729 A; 1095 C; 975 G; 363 T; 0 U; 0 Other;
```

Qy  Best Match

```
Query Match             84.5%;  Score 1275.2;  DB 9;  Length 3162;
Best Local Similarity   92.5%;  Pred. No. 1.6e-148;
Matches 1398;  Conservative    0;  Mismatches  78;  Indels  36;  Gaps  4;
```

```
Qy      1 ATGGGCGCCCGGGCCCAGCATCCTGCGCGGCGACGAAGCTCGGACAAGAGTGGGAGAGATCCGC 60
          ||||||||||||||||||||||||||||||||||||||||
Db      7 ATGGGCGCCCGGGCCCAGCATCCTGCGCGGCGACGAAGCTCGGACAAGAGTGGGAGAGATCCGC 66

Qy     61 CTGCGCCCCGGGCGGGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
          ||||||||||||||||||||||||||||||||||||||||
Db     67 CTGCGCCCCGGGCGGGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 126
```

Db 1144 AAGGGCACCCCGCATCGTCAAGTCTTCAACTGCGCGCAAGGAGGAGGGCCACATCGCCCGC 1203

Qy 1198 AACTGCGCGCGCCCCCCGCCCCCAAGGAGGGCTGCTGCGAAGTGCGGCGAAGGAGGCCCAAGGAGGGCCCAGATG 1257

Db 1204 AACTGCGCGCGCCCCCCGCCCCCAAGGAGAAGGGCTGCTGCGGAAGTGCGGCGAAGGAGGCCCACCAGATG 1263

Qy 1258 AAGGACTGCACGAGCGCCAACTTCCTGCCCAAGATCTGGGCCCACAACGGC 1317

Db 1264 AAGGACTGCACGAGCGCCAACTTCCTGGGCCAAGATCTGGGCCCACAACGGC 1323

Qy 1318 CGCCCCGGCCCAGCCAACCCACCGGCACCGGCACCAAGGCCCACCGGCC 1377

Db 1324 CGCCCCGGCCAACTTCCTGCACAGAGCCG----------CCCGACCCCACCGCC 1365

Qy 1378 CCCCCCGCCCGAGAGCTTCCGGTTCGGGGAGGAGAGACCACCCCCCGGCACCCCGCCCCCCCAAG 1437

Db 1366 CCCCCCGCCCGAGAGCTTCCGCTTCGGGGAGGAGACCACCCCCGCCAGCGCAAG 1425

Qy 1438 GACCCGCGAGCGCCTTACCGCGGCCCCCCCTACGGCGCACCCTTGTTCGGGCGAGGCGGCGCCC 1497

Db 1426 G------------ACCGCGAGACCCTGACCAGCCCTGAGAGAGCCCTGTTCGGGCAACGAGACCCC 1473

Qy 1498 CTGAGCCAGTAA 1509

Db 1474 CTGAGCCAGAAA 1485

RSSULT 13

ID AD013237 standard; DNA; 3462 BP.
XX
AC AD013237;
XX
DT 18-DBC-2003 (first entry)
XX
DE DNA of HIV construct GagProtInaRtmut_C SEQ ID NO 16.
XX
KW expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS Human immunodeficiency virus.
XX
XX
PN WO2003004620-A2.
XX
PD 16-JAN-2003.
XX
PF 05-JUL-2002; 2002WO-US021420.
XX
PR 05-JUL-2001; 2001US-0303192P.
PR 31-AUG-2001; 2001US-0316860P.
PR 16-JAN-2002; 2002US-0348871P.
XX
PA (CHIR ) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX
PI Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR WPI; 2003-221593/21.
XX
PT New expression cassette comprising a polynucleotide sequence encoding a
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT Prot, or Rev polypeptide, useful for immunization, or generating
PT packaging cell lines;
XX
PS Disclosure; Fig 13; 301pp; English.
XX
CC The invention relates to a novel expression cassette comprising a
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC expression cassette can be used to treat HIV type C by gene therapy or
CC used in the development of a vaccine. The gene delivery vector is
CC administered intramuscularly, intranasally,
CC subcutaneously, intradermally, transdermally, intravaginally,

CC intrarectally, orally or intravenously. The expression cassette is useful
CC for immunisation, generating packaging cell lines and producing HIV
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC Type C related sequence of the invention.
XX
SQ Sequence 3462 BP; 790 A; 1198 C; 1068 G; 406 T; 0 U; 0 Other;

Query Match       84.5%;   Score 1275.2;  DB 9;   Length 3462;
Best Local Similarity   92.5%;   Pred. No.1.6e-148;
Matches 1398;   Conservative   0;   Mismatches 78;   Indels 36;   Gaps 4;

Qy 1 ATGGGAGCAGCTGCAGCGCCAGCCATCCTGCCGGGCGGCGGAGGTCCGACCAAGGCC 60
Db 1 ATGGGGCGGCGCCGGCGGCCAGCATCCTGCGCGGGCGGCGAGGAGGTCGGACCAAGGAGGCC 60

Qy 61 CTCGCGAC---GGCAAGGTGAGCCAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCGAGATGGG 120
Db 67 CTCGCGCCCCGGGCGGCAAGGTGAGCCAGAAGCCAGAACTACCCCATCGTGCAGAGGACCTGGCCAGAGAG 126

Qy 121 CTGGGAGAGGGCTTGCCCCTGAACCCCGGCGCCTGCTGCTGGACCCGGCAGCGAGATC 180
Db 127 CTGGGAGAGTTCCCCCTGAACCCCGGCGCCTGCTGGACGAGCAGCGAGATC 186

Qy 181 ATGAAGCAGCTGCAGCCCCGGCCCCTGCTGCACCCTGTACTGGCGTGCACGCCGGCATCGAGGTCCGACCAAGGCC 240
Db 187 ATCCGCCCAGCTGCACCCCGGCCCTGCACCCGGGCGAGGAGCGCCTGTTCAAC 246

Qy 241 ACCGTGGCCACCCTGTACTGGCGTGCACGCCGGGCATCGAGGTCCGACCAAGGCC 300
Db 247 ACCGTGGGCCACCCTGTACTGCGTTGCACGCGACGTCGGAGGTCCGCGGCGACACCAAGGAGGAGGCC 306

Qy 301 CTGGACAAGATGGAGGAGGCAGACAGCAAGTCCCAGGCAGGAAGATCGAGGAG 360
Db 307 CTGGACAAGATCGAGGACGAGGAGAAACAAGTGCCACAGGAACCAGCAGGCAGAGGCC 366

Qy 361 GCCGAC----GGCAAGGTGAGCCAGAACTACATCGTCAGAACCTGCAGAGATG 417
Db 367 GCCGACAAGGGCAAGGTGAGCCAGCAGACCACCCCATCGTCAGAACCTGCAGAGATG 426

Qy 418 GTGCACAGGACCATCAGCCCCCGCACCCTGAACGCCTGGGTCAAGGTGATCGAGGAG 477
Db 427 GTGCACACCAGGGCCATCAGCCCCCTGAACGCCTGGGTGAAGGTGATCGAGGAAG 486

Qy 478 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACGGCCCTGAGCGAGGGCGCCACCCCAG 537
Db 487 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACGGCCCTGAGCGAGGGCGCCACCCAG 546

Qy 538 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 597
Db 547 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 606

Qy 598 GACACCATCAACGAGGAGGCAGCGCCTGCACCCGTGCAGCGGCGGCCC 657
Db 607 GACACCATCAACGAGGAGGCAGCGCCTGCACCCGTGCAGCGGCGGCCC 666

Qy 658 GTGGGCCCCCGGCCCAGATGCGGGACCCCCCCGCCGAGCATCGCCACCACCAGCACC 717
Db 667 ATCGCCCCCCGGCCAGATGCGGGACCCCCCCGCCGAGCATCGCCACCACCAGCACC 726

Qy 718 CTGGCAGGAGCAGATGCGCCTGGAGTGAACGACGAACCAACCCCGTGCCGGGACATCTAC 777
Db 727 CTGCAGGAGCAGATGCGCCTGAGTGAACGACCAGCAGCATCTAC 786

Qy 778 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCGTGAGCATC 837
Db 787 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCGTGAGCATC 846

Qy 838 CTGGACATCCGGCAGGGACCCTTCGCGACTACGTGGACCGCTTCTTCAAG 897
Db 847 CTGGACATCAGCAGGGACCCCTTCGCGACTACGTGGACCGCTTCTTCAAG 906

Qy 898 ACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTG 957
Db 907 ACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTG 966

RESULT 14

ACA03523

ID   ACA03523 standard; DNA; 4419 BP.
XX
AC   ACA03523;
XX
DT   22-MAY-2003  (first entry)
XX
DE   Synthetic DNA encoding immunogenic HIV peptide #6.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; de.
XX
OS   Synthetic.
XX
PN   WO2003004657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349760P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
PS   Example 1; Fig 11; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotides or vaccines for generating humoral or
CC   cellular immune responses against HIV in a subject, or for inducing
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene encoded proteins are useful for generating
CC   polynucleotide is also useful for DNA immunisation of, or for generating
CC   an immune response (e.g. a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
CC   human immunodeficiency virus immunogenic peptide
XX
SQ   Sequence 4419 BP; 980 A; 1582 C; 1371 G; 486 T; 0 U; 0 Other;
XX

Query Match    84.5%;  Score 1275.2;  DB 7;  1371 G;  486 T;  0 U;  0 Other;
Best Local Similarity   92.5%;  Pred. No. 1.5e-148;
Matches 1398;  Conservative   0;  Mismatches  78;  Indels  36;  Gaps  4;

KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
OS   Human immunodeficiency virus.
XX
PN   WO200304620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 16; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 4419 BP; 1582 C; 1371 G; 486 T; 0 U; 0 Other;

```
Query Match             84.5%;  Score 1275.2;  DB 9;  Length 4419;
Best Local Similarity   92.5%;  Pred No. 1.5e-148;
Matches  1398;  Conservative    0;  Mismatches   78;  Indels   36;  Gaps    4;

Qy     1 ATGGGGCGCGCCGGGGCCAGCATCCTNCGCGGCGGGCGAGAAGCTTGGACAAGTGGGAAGATCCGC 60
           |||||||||||||||||||||  ||||||||||||||||
Db     7 ATGGGGCGCCCGGGGCCAGCATCCTGCGGCGGG                                 66

Qy    61 CTGGCCCCCGCGGCGGAGAGCACTACATCGTGTGGGGCCAGCCCGAG 120
           ||||||||||||||||
Db    67 CTGCGCGGCGGCGAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCCGGAG 126

Qy   121 CTGGAGGGGCTTCGCGTCGACCTGGCAGCCGGGCATCCGGC 180
           |||||||||||
Db   127 CTGGAGAAGTTCGGCCCTGAACCCGGCCTGCTGGAGACCAGCGAGGGGCTGAACCAGATC 186

Qy   181 ATGAAGCAGCTGCAGGACCACTACATGCTGAGGCCGGAGGCCCGAGATC 240
           || |||||||||||||||||
Db   187 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCACCGAGGAGCTGAGGCCGCTGTTCAAC 246

Qy   241 ACCGTGGCCCACCCTGTACTGGTGCAGCGCCGGCATCGAGGTCCCGGACATC 300
           |||||||||||||||||| |
Db   247 ACCGTGGCCCACCCTGTACTGGTGCACGCAGGATCGAGGTCCCGGACACCAAGGAGGCC 306

Qy   301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAGCAGGAGAACAAGGCGCCAAGGAG 360
           ||||||||||||||||||| |||||||||| |||
Db   307 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGAAGAGATCCAGAGAAGGCGGAGGCC 366

Qy   361 GCCGGAC--GGCAAGGTGAGCCGAACTACCCCATCGTTCAGAACCTTCGCAGGGCCCAGATG 417
           ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||
Db   367 GCCCACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTACGCAGGGCCCAGATG 426
```
```
Qy   658 GTGGGCCCCCGCGGCGGGATNGCGGCGGACCCCCGGCGGCAGCGATCGCGGGGCCCACCAGCACC 717
           ||||||||||||||||
Db   667 ATCGGCCCCCGCGGCGGGATNGCGGCGGACCCCGGCCGGCAGCATGCGGGGGCCCACCAGCACC 726

Qy   718 CTGCAGGAGCAGATCGCCTGGATGACGAGCAACCCCCGGCCGGCAGCATGCGGGGGCACATCTAC 777
           ||||||||||||||||
Db   727 CTGCAGGAGCAGATCGCCTGGATGACGAGCAACCCCCCGGTGCCCGTGGGGCGGCGGGGACATCTAC 786

Qy   778 AAGGCCGTGGATCATCTGGGGCTGCTGAACACAGATCGTCTGGGGATGTACAGCCCGTGAGCGATC 837
           |||||||||||||||
Db   787 AAGCCGGTGGATCATCTGGGGCTGACGAGCAACCCCCCGGTGGGATTAGACAGCCCCTGAGCGATC 846

Qy   838 CTGGACATCCCGCAGGGGCCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 897
           |||||||||||||||
Db   847 CTGGACATCAAGCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 906

Qy   898 ACCCTGCGCGCCGAGCAGGCAACCCAGGATGTGAAGAACTGGATGACCGAGACCCTGCTG 957
           |||||||||||||||
Db   907 ACCCTGCGCGCCGAGCAGGCAACCCAGGATGTGAAGAACTGGATGACCGAGACCCTGCTG 966

Qy   958 GTGCAGAACGCCAACCCGGACTGCAAGACCATCCTGCGGGCCCTGGGGCCAGGCCCCACC 1017
           |||||||||||||||
Db   967 GTGCAGAACGCCAACCCGGACTGCAAGACCATCCTGCGGGCCCTGGGGCCAGGCCCCAGC 1026

Qy  1018 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGGCGGCCCGGGCACAAGGCCCGCGTG 1077
           |||||||||||||||
Db  1027 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGGCGGCCCGGGCACAAGGCCCGCGTG 1086

Qy  1078 CTGGCCGAGGCCATGAGCCAGGCCAACATCATGATGCAGAAGAGCAACTTC 1137
           |||||||||||||  |
Db  1087 CTGGCCGAGGCCATGAGCCAGGCCAACCACC--AGCCTGATGATGCAGAAGAGCAACTTC 1143

Qy  1138 AAGGCGCCCCCCCGGCCGCACGTCAAGTGCTTCAACTGCGGCAAGGAGGGGCACATCGCCAAG 1197
           N|||||||||||||||
Db  1144 AAGGGCCCCCCGGCCGCACGTCAAGTGCTTCAACTGCGGCAAGGAGGGGCACATCGCCAAG 1203

Qy  1198 AACTGCGCGCGCGCCCCGGCAAGAACGGGCTGCTTCGGAAGTGCGGCAAGGAGGGGCACCAGATG 1257
           |||||||||||||||
Db  1204 AACTGCGCGCGCGCCCCCGGCAAGAAGGGGTGCTGCGGAAGTGCGGCAAGGAGGGGCACCAGATG 1263

Qy  1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCCAGCCCAACAAGGGC 1317
           |||||||||||||||
Db  1264 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCAGCCCAACAAGGGC 1323

Qy  1318 CGCCCCGGCAACTTCCTGCAGAACCCCAGCGCCCCGCCCCCGCGCCCCACCGTGCCCACCGCC 1377
           ||||||||||||      |||
Db  1324 CGCCCCGGCAACTTCCTGCAGAGCCGG------CCCGAGCCCACCGCCCC 1365

Qy  1378 CCCCCGCCGCCGAGAGCTTCGAGGAGACCACACCCGCGCCCAAGCAGGAGCCAGCCCGAG 1437
           |||||||||||||||
Db  1366 CCCCCCGCCGCCGAGAGCTTCGAGGAGACCACACCCGCGCCCAAGCAGGAGCCAGCCCGAG 1425

Qy  1438 GACCGGAGCCCTACCCGCCGCCCTGCCCTGGAGAGCCTGGGGCAGGGCGGGGGGCCCC 1497
           |||||||       |||
Db  1426 G-------ACCCGGAGACCCTGAACAGCCTGAGAGCCTGTTCGGGCCAGAGACCCGGGATC 1473

Qy  1498 CTGGAGCCAGTAA 1509
           |||||||||
Db  1474 CTGAGCCAGAAA 1485
```

Qy  1498 CTGAGCCAGTAA 1509
         |||||||| || ||
Db  1474 CTGAGCCCAGAAA 1485

Qy   418 GTGCACCAGGCCATCAGCCCCCCACCCTGAACGCCTGGGGTGAAGGTGATCGAGGAGAAG 477
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427 GTGCACCAGGCCATCAGCCCCCCACCCTGAACGCCTGGGGTGAAGGTGATCGAGGAGAAG 486

Qy   478 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCGCCACCCCCAG 537
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGCGCCACCCCCAG 546

Qy   538 GACCTGAACACGATGTTGAACACCGTGGGGCGCCACCAGCGCGCCCATGCAGATGCTGAAG 597
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547 GACCTGAACACGATGTTGAACACCGTGGGGCGCCACCAGCGCGCCCATGCAGATGCTGAAG 606

Qy   598 GACACCATCAACGAGGAGGAGGCCGGCCGAGTGGGACCGGCCTTGCACCCCGTGCAGGGCCCC 657
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607 GACACCATCAACGAGGAGGAGGCCGGCCGAGTGGGACCGGCCTTGCACCCCGTGCAGGGCCCC 666

Qy   658 GTGGGCCCCCGGCCAGATCGCGACACCCGGCCAGCGCGACTCGCGCGGCGCCCACCCACC 717
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667 ATCGCCCCCGGCCAGATCGCGACACCCGGCCAGCGCGACATCGCCGGCGCCCACCCACC 726

Qy   718 CTGGCAGGAGCAGATGCGCCTGGATGACCAGCAGCACCCCCGTCCCGGCGGCACATCTAC 777
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727 CTGGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCGTCCCGGCGGGCGACATCTAC 786

Qy   778 AAGCGGGTGGATCATCCTGGGCCTGAAACAAGATGTGCGGAGTGTACAGCCGTGAGCATC 837
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   787 AAGCGGGTGGATCATCCTGGGCCTGAAACAAGATCGTGCGGAGTGTACAGCCGTGAGCATC 846

Qy   838 CTGGACATTCCGCCAGGCGCCCCAAGGAGCCCTTCGCGACTACCTGGCCGCTTCTTCAAG 897
         |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   847 CTGGACATCAAGCAGCGGCCCCCAAGGAGCCCTTCGCGGCTACCTGGCACCGCTTCTTCAAG 906

Qy   898 ACCCTGCGGCGCGGAGCGCAGGGCCACCCAGCACGGCGTGAAGAACTGGATGACCCTGCTG 957
         | || |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   907 ACCCTGCGGCGCCCAGCACGAGGTGAAAGAACTGGATGACCGACACCCCTGCTG 966

Qy   958 GTGCAGGAACGGCCAACCCGGACTCGCAAGACCATCCTGCGCGCTTCTTGGCGCCCGGCGCCACC 1017
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   967 GTGCAGGAACGGCCAACCCGGACTCGCAAGACCATCCTGCGCGCTTCTTCGGCGCCCGGCGCCAGC 1026

Qy  1018 CTGGGAGGAGATGATGACCGGCCTCGCCAGGGGTGGGGCGGCGCCCGGGCCGGCGTG 1077
         ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  1027 CTGGAGGAGATGATGACCGGCCTTGCCAGGGGTGGGGCGGCCCCACCCACAAGGGCCGCGGTG 1086

Qy  1078 CTGGGCCGGAGGCGGATGAGCGGCCCAACAGCGTGAACATCATCATGTGCGGAAGAGCACAACTTC 1137
         |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  1087 CTGGCCCGGAGGCGGATGAGCGGCCAACACC----AGCGTGATGATGCAGAAGAGCACAACTTC 1143

Qy  1138 AAGGGCCCCCGGCGGCGCAAGTCAAGTGCTTCAACTGGCCCAAGGCCGGCCACATCGCCAAG 1197
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1144 AAGGGCCCCGGCGGCGCAAGTGCTTCAACTGGCCCAAGGCCGGCCCACATCGCCAAG 1203

Qy  1198 AACTGCGCGCGCCCCCCCCGCAAGGAAGGGCTGCTCGAAGTCGGCGGCAAGGAGGGCCGCCACCAGATG 1257
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1204 AACTGCGCGCGCCCCCGCAGAGAAGGGCTGCTCGAAGTCGGCGCAAGGAGGGCCACCAGATG 1263

Qy  1258 AAGGACTTGCACCGGAGCGCCAGGCCGCCAACTTCCTGGGCAAGATCTGGCCCCAGCCAAAGGC 1317
         ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db  1264 AAGGACTGCACCGGAGCGCCAGGCCCACAACTTCCTGGGCAAGATCTGGCCCCAGCACAAGGGC 1323

Qy  1318 CGGCCCCGGCCAACTTCCTGCAGAACGGCCAGCCGGAGCCCGGCCCGGTGCCCACCGC 1377
         |||||||||||||||||| |||                    ||||||||||||||||
Db  1324 CGGCCCCGGCCAACTTCCTCCGCACAGACGCG--------CCCCGGAGCCCCACACCCC 1365

Qy  1378 CCCCCGCCCCGAGAGCTTCCGCTTCGCGCTTCGAGAGGAGAACCCCGCCCCCAAGCCAGCAGACCCCAAG 1437
         |||||||||||||||||||||||||||||||||||||| || ||||||||||||
Db  1366 CCCCCCCGCCGAGAGCTTCCGCTTCGAGGACCACCCCCCCGGCGAGAGCGGCAGCAAG 1425

Qy  1438 GACGCGGAGCCCTTACCCGGAGCCCCTGACCCGCGCAGCCTGTTCGGGCAGCGGGGCCCC 1497
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1426 G--------ACCGCGAGACCCTGACCAGCCTGTGAAGAGCGGCTGTTCGGGCAACGAGACCCC 1473

QM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen ltd.

Run on:    June 1, 2004, 10:26:25 ; Search time 5848 Seconds
    (without alignments)
    11184.104 Million cell updates/sec

Title:    US-09-475-704A-4
Perfect score:    1509
Sequence:    1 atgggcgccgccgccagcat..........gcggcccctgagccagtaa 1509
Scoring table:    IDENTITY_NUC
Scoring:    Gapop 10.0 , Gapext 1.0
Searched:    3470272 seqs, 2167151699S residues

Total number of hits satisfying chosen parameters:    6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
    Maximum Match 100%
    Listing first 45 summaries

Database :    GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_vrt:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1509 | 100.0 | 1509 | 6 | AX455888 | AX455888 Sequence |
| 2 | 1507.4 | 99.9 | 1509 | 6 | AX468544 | AX468544 Sequence |
| 3 | 1504.2 | 99.7 | 1509 | 6 | AX468548 | AX468548 Sequence |
| 4 | 1489.8 | 98.7 | 1509 | 6 | AX455905 | AX455905 Sequence |
| 5 | 1297.6 | 86.0 | 1494 | 6 | AX455935 | AX455935 Sequence |
| 6 | 1294.8 | 85.8 | 1491 | 6 | AX455983 | AX455983 Sequence |
| 7 | 1291.4 | 85.6 | 1485 | 12 | AY181195 | AY181195 Synthetic |
| 8 | 1276.8 | 84.6 | 1479 | 6 | AX455887 | AX455887 Sequence |
| 9 | 1276.8 | 84.6 | 1479 | 6 | AX468543 | AX468543 Sequence |
| 10 | 1271.2 | 84.2 | 1479 | 6 | AX468547 | AX468547 Sequence |
| 11 | 1260.8 | 83.6 | 1479 | 6 | AX455904 | AX455904 Sequence |
| 12 | 1202.2 | 79.7 | 4288 | 6 | AX149648 | AX149648 Sequence |
| 13 | 1201 | 79.6 | 1509 | 12 | AF201927 | AF201927 Synthetic |
| 14 | 1201 | 79.6 | 1515 | 6 | BD263637 | BD263637 Improved |
| 15 | 1201 | 79.6 | 1515 | 6 | AR373320 | AR373320 Sequence |
| 16 | 1201 | 79.6 | 4472 | 6 | BD263697 | BD263697 Improved |
| 17 | 1201 | 79.6 | 4472 | 6 | AR373380 | AR373380 Sequence |
| 18 | 1201 | 79.6 | 4608 | 6 | BD263698 | BD263698 Improved |
| 19 | 1201 | 79.6 | 4608 | 6 | AR373381 | AR373381 Sequence |
| 20 | 1201 | 79.6 | 4689 | 6 | BD263696 | BD263696 Improved |
| 21 | 1201 | 79.6 | 4689 | 6 | AR373379 | AR373379 Sequence |
| 22 | 1201 | 79.6 | 4766 | 6 | BD263695 | BD263695 Improved |
| 23 | 1201 | 79.6 | 4766 | 6 | AR373378 | AR373378 Sequence |
| 24 | 1199 | 79.5 | 2031 | 6 | BD263640 | BD263640 Improved |
| 25 | 1199 | 79.5 | 2031 | 6 | AR373323 | AR373323 Sequence |
| 26 | 1186 | 78.6 | 9166 | 6 | AX427930 | AX427930 Sequence |
| 27 | 1170.8 | 77.6 | 9788 | 6 | AX427936 | AX427936 Sequence |
| 28 | 1145 | 75.9 | 1539 | 6 | AX306431 | AX306431 Sequence |
| 29 | 1143.2 | 75.8 | 1548 | 12 | AF287354 | AF287354 Synthetic |
| 30 | 1119 | 74.2 | 6438 | 6 | AX427919 | AX427919 Sequence |
| 31 | 1119 | 74.2 | 9194 | 6 | AX427925 | AX427925 Sequence |
| 32 | 1119 | 74.2 | 9194 | 6 | AX427926 | AX427926 Sequence |
| 33 | 1119 | 74.2 | 12411 | 6 | AX427927 | AX427927 Sequence |
| 34 | 1117.8 | 74.1 | 1847 | 12 | AF202464 | AF202464 Synthetic |
| 35 | 1117.8 | 74.1 | 1847 | 12 | AF202465 | AF202465 Synthetic |
| 36 | 1117.8 | 74.1 | 1853 | 6 | BD263638 | BD263638 Improved |
| 37 | 1117.8 | 74.1 | 1853 | 6 | AR373321 | AR373321 Sequence |
| 38 | 1117.8 | 74.1 | 1865 | 6 | BD263700 | BD263700 Improved |
| 39 | 1117.8 | 74.1 | 1865 | 6 | BD263701 | BD263701 Improved |
| 40 | 1117.8 | 74.1 | 1865 | 6 | AR373383 | AR373383 Sequence |
| 41 | 1117.8 | 74.1 | 1865 | 6 | AR373384 | AR373384 Sequence |
| 42 | 1117.8 | 74.1 | 4319 | 6 | BD263639 | BD263639 Improved |
| 43 | 1117.8 | 74.1 | 4319 | 6 | AR373322 | AR373322 Sequence |
| 44 | 1103.8 | 73.1 | 9170 | 6 | AX427928 | AX427928 Sequence |
| 45 | 1103.8 | 73.1 | 9189 | 6 | AX427921 | AX427921 Sequence |

ALIGNMENTS

RESULT 1
AX455888
LOCUS    AX455888    1509 bp    DNA    linear    PAT 06-JUL-2002
DEFINITION    Sequence 4 from Patent WO0204493.
ACCESSION    AX455888
VERSION    AX455888.1    GI:21714801
KEYWORDS    .
SOURCE    synthetic construct
  ORGANISM    synthetic construct
    artificial sequences.
REFERENCE    1
  AUTHORS    zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.
  TITLE    Polynucleotides encoding antigenic hiv type c polypeptides,
    polypeptides and uses thereof
  JOURNAL    Patent: WO 0204493-A 4 17-JAN-2002;

CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

```
FEATURES
    source          Location/Qualifiers
                    1..1509
                    /organism="synthetic construct"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32630"
                    /note="synthetic Gag of HIV strain AP110967"

ORIGIN
    Query Match      100.0%;  Score 1509;  DB 6;  Length 1509;
    Best Local Similarity  100.0%;  Pred. No. 1.3e-150;
    Matches 1509;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy       1 ATGGGCGCCCCGGCCCGCCCAGCATCCTGCGCGCGGGCGAGGAGGTGGACAAGGTGGGAGAAGATCCGC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGGCGCCCCGGCCCGCCCAGCATCCTGCGCGCGGGCGAGGAGGTGGACAAGGTGGGAGAAGATCCGC 60

Qy      61 CTGCGCCCCGGCGGCGGGAGCAAGAAGAGCTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTGCGCCCCGGCGGCGGGAGCAAGAAGAGCTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120

Qy     121 CTGGGAGGGCCTTGCCCTTGAAACCCCGGGCCTGCTTGGAGACCGCCGGAGCCCGGAGGGCGCAAGCAGAATC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTGGGAGGGCCTTGCCCTTGAAACCCCGGGCCTGCTTGGAGACCGCCGGAGCCCGGAGGGCGCAAGCAGATC 180

Qy     181 ATGAAGCAGCTCCAGCCCTGCCACGCCCGGCCACCGGAGCCGGAGCCCCGGAGGCCTGCTGTACAAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATGAAGCAGCTCCAGCCCTGCCACGCCCGGCCACCGGAGCCGGAGCCCCGGAGGCCTGCTGTACAAC 240

Qy     241 ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTCCGGCGACCACCAAGGAGGAGCC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTCCGGCGACCACCAAGGAGGAGCC 300

Qy     301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAAACCCAGCAGGCCGCCGCGGACCCAAGGAG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAAACCCAGCAGGCCGCCGCGGACCCAAGGAG 360

Qy     361 GCCGACACCGGCAAGGTGAGCCCCAAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GCCGACACCGGCAAGGTGAGCCCCAAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTG 420

Qy     421 CACCAGGCCCATCAGCCCCCTGAAACGCCCACCCATGGGGTTGATCGAGGGCAGGAAGCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CACCAGGCCCATCAGCCCCCTGAAACGCCCACCCATGGGGTTGATCGAGGGCAGGAAGCC 480

Qy     481 TTCAGCCCCGAGGTGATCCCCATGTTCACGGCCCTGAGCGAGGGCGCCACCCCCCAGGAC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TTCAGCCCCGAGGTGATCCCCATGTTCACGGCCCTGAGCGAGGGCGCCACCCCCCAGGAC 540

Qy     541 CTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGAC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGAC 600

Qy     601 ACCATCAACGAGGAGGCGGCCGCCGAGTGGGACCGCGTGCACCCGGTGCACGCCGGCCCCGTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ACCATCAACGAGGAGGCGGCCGCCGAGTGGGACCGCGTGCACCCGGTGCACGCCGGCCCCGTG 660

Qy     661 GCCCCCGGCCAGATGCGGGACCACCCGCGGCCGCCGACATCGGCGGCACCCCGGGCGAGGCCCCCGTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GCCCCCGGCCAGATGCGGGACCACCCGCGGCCGCCGACATCGGCGGCACCCCGGGCGAGGCCCCCGTG 720

Qy     721 CAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCGTGCCCGTGGGCGACAATCTACAAG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCGTGCCCGTGGGCGACAATCTACAAG 780

Qy     781 CGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTG 840

Qy     841 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 900

Qy     901 CTGCGCGCCGAGCAGGCCAGCCCAGCCCAGGGACGCTGAAGAACTTGGATGACCGGAGCCCCGTGCTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTGCGCGCCGAGCAGGCCAGCCCAGCCCAGGGACGCTGAAGAACTTGGATGACCGGAGCCCCGTGCTG 960

Qy     961 CAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCGGCCGCCACCCTG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 CAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCGGCCGCCACCCTG 1020

Qy    1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCGGGCCACAAGGCCCGCGTG 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCGGGCCACAAGGCCCGCGTG 1080

Qy    1081 GCCGGAGGCGGATGAGCCAGGCCCAGACAGCCGCCAACATCATGATGCAGAAGAGCAACTTCAAG 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GCCGGAGGCGGATGAGCCAGGCCCAGACAGCCGCCAACATCATGATGCAGAAGAGCAACTTCAAG 1140

Qy    1141 GGCCCCCGGCCGCCAAGGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACACTGCCAAGAAC 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GGCCCCCGGCCGCCAAGGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAAC 1200

Qy    1201 TGCCGCGCCCCGCCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TGCCGCGCCCCGCCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260

Qy    1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCAGCCCCAAGGGGCGC 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCAGCCCCAAGGGGCGC 1320

Qy    1321 CCCGCGCGAGAGCTTCGCTTCGGCGAAGAACCCAGCGAGCGCGAGCCCACCGTGCCCCAAC 1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CCCGCGCGAGAGCTTCGCTTCGGCGAAGAACCCAGCGAGCGCGAGCCCACCGTGCCCCAAC 1380

Qy    1381 CCCCGGCCGAGAGCGGGCCCAGCTCCAGAGGAGCACCACCCCCCAGAAGCAGGAGCCCAAGGAC 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 CCCCGGCCGAGAGCGGGCCCAGCTCCAGAGGAGCACCACCCCCCAGAAGCAGGAGCCCAAGGAC 1440

Qy    1441 CGCGAGCCCTACCCCGAGCCCCCTGACCCCGGCCTGTTCGGGCAGGCGGCCCCCCTG 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 CGCGAGCCCTACCCCGAGCCCCCTGACCCCGGCCTGTTCGGGCAGGCGGCCCCCCTG 1500

Qy    1501 AGCCAGTAA 1509
           |||||||||
Db    1501 AGCCAGTAA 1509
```

```
RESULT 2
AX468544

LOCUS       AX468544                1509 bp    DNA     linear   PAT 16-JUL-2002
DEFINITION  Sequence 64 from Patent WO0226209.
ACCESSION   AX468544
VERSION     AX468544.1  GI:21901374
KEYWORDS
SOURCE      Human immunodeficiency virus 1 (HIV-1)
  ORGANISM  Human immunodeficiency virus 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1
  AUTHORS   O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
            Singh,M., Ulmer,J. and Dubensky,T.W.
  TITLE     Microparticles for delivery of the heterologous nucleic acids
  JOURNAL   Patent: WO 0226209-A 64 04-APR-2002;
            CHIRON CORPORATION (US)
FEATURES             Location/Qualifiers
    source          1..1509
                    /organism="Human immunodeficiency virus 1"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:11676"
ORIGIN
```

```
    Query Match      99.9%;  Score 1507.4;  DB 6;  Length 1509;
    Best Local Similarity  99.9%;  Pred. No. 1.9e-150;
    Matches 1508;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;
```

```
Qy      1 ATGGGCGCCCCCGCGCCCAGCATCCTGCGCGGGCGAGAAGCTGGACAAGTGGGAGAAGATCCGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGGCGCCCCCGCGCCCAGCATCCTGCGCGGGCGAGAAGCTGGACAAGTGGGAGAAGATCCGC 60
```

RESULT 3

| LOCUS | AX468548 | 1509 bp | DNA | linear | PAT 16-JUL-2002 |
|---|---|---|---|---|---|

DEFINITION  Sequence 68 from Patent WO0226209.
ACCESSION   AX468548
VERSION     AX468548.1  GI:21901378
KEYWORDS    .
SOURCE      Human immunodeficiency virus 1 (HIV-1)
  ORGANISM  Human immunodeficiency virus 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1
  AUTHORS   O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
            Singh,M., Ulmer,J. and Dubensky,T.W.
  TITLE     Microparticles for delivery of the heterologous nucleic acids
  JOURNAL   Patent: WO 0226209-A 68 04-APR-2002;
            CHIRON CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1509
                     /organism="Human immunodeficiency virus 1"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11676"
ORIGIN

Query Match             99.7%;  Score 1504.2;  DB 6;  Length 1509;
Best Local Similarity   99.2%;  Pred. No. 4.1e-150;
Matches 1497;  Conservative 12;  Mismatches 0;  Indels 0;  Gaps 0;

RESULT 4

```
LOCUS       AX455905                1509 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 21 from Patent WO0204493.
ACCESSION   AX455905
VERSION     AX455905.1  GI:21714897
KEYWORDS
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 21 17-JAN-2002;
            CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..1509
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="synthetic Gag coding sequence of HIV strain
                     AF110967"
ORIGIN

Query Match     98.7%;  Score 1489.8;  DB 6;  Length 1509;
Best Local Similarity   99.2%;
Matches 1497;  Conservative   0;  Mismatches 12;  Indels  0;  Gaps  0;
```

Qy        1 ATGGGCGCCCCGCCGGCCAGCATCCTGCCCGGCCAGCAACTGCACAGGTGGGAAGAATCCGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGGGCGCCCCGCCGGCCAGCATCCTGCCCGGCCAGCAACTGCACAGGTGGGAAGAATCCGC 60

Qy       61 CTGCGCGGCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CTGCGCGGCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120

Qy      121 CTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCAAGCAGATC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCAAGCAGATC 180

Qy      181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCTCTACAAC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCTCTACAAC 240

Qy      241 ACCGTGGCCACCCTGTACTGTGCGTGGACACCGGGGCCATCGCCGTGACTGGAGGC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 ACCGTGGCCACCCTGTACTGTGCGTGGACACCGGGGCCATCGCCGTGACTGGAGGC 300

Qy      301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACGCAGCAGGCCAAGGAG 360
            ||::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 CTGGACAAGATCGAGGAGGAGCAGAACAAGWCCCAGCAGAAGACGCAGCAGGCCAAGGAG 360

Qy      361 GCCGGACAGGCGAAGTGAGCCAGGAGCTACCACCGCCGCCAGCAGATGGTG 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GCCGGACAGGCGAAGTGAGCCAGGAGCTACCACCGCCGCCAGCAGATGGTG 420

Qy      421 CACCAGGCCCATCAGCCCTGAACGCCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAAGCC 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 CACCAGGCCCATCAGCCCTGAACGCCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAAGCC 480

Qy      481 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGAC 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGAC 540

Qy      541 CTGAACACGATGTTGAACACCGTGGGGGGCGCCACCAGGCGCCCCATGCAGATGCTGAAGGAC 600
            |||::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 CTGAACACSATGTTGAACACCGTGGGGGGCGCCACCAGGCGCCCCATGCAGATGCTGAAGGAC 600

Qy      601 ACCATCAACGAGGAGGCGGCCGAGTGGGACCGCCTGCACCCCGTGCAGCCCGGCCCCGTG 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 ACCATCAACGAGGAGGCGGCCGAGTGGGACCGCCTGCACCCCGTGCAGCCCGGCCCCGTG 660

Qy      661 GCCCCCGGCCAGATGCGGGACCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACCCTG 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GCCCCCGGCCAGATGCGGGACCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACCCTG 720

Qy      721 CAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCGGTGCCCGTGGGCGACATCTACAAG 780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 CAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCGGTGCCCGTGGGCGACATCTACAAG 780

Qy      781 CGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGSATGTACAGCCCCGTGAGCATCCTG 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 CGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGSATGTACAGCCCCGTGAGCATCCTG 840

Qy      841 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 900

Qy      901 CTGCGCGCCGAGCAGGCCACCCAGGAAGTGAAGAACTGGATGACCGAGACCCTGCTGGTG 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 CTGCGCGCCGAGCAGGCCACCCAGGAAGTGAAGAACTGGATGACCGAGACCCTGCTGGTG 960

Qy      961 CAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCGCTCTCGGCCCCGCCCTG 1020
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 CAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCGCTCTCGGCCCCGCCCTG 1020

Qy     1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTGCTG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTGCTG 1080

Qy     1081 GCCGAGGCCATGAGCCAGGTGAACAGCACCGCCATCCTGATGCAGCGCAGCAACTTCAAG 1140
            ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 GCCGAGGCSATGAGCCAGGTGAACAGCACCGCCATCCTGATGCAGCGCAGCAACTTCAAG 1140

Qy     1141 GGCCCCAAGCGCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAACTG 1200
            ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 GGCCCCAAGCGCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAACTG 1200

Qy     1201 TGCCGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTG 1260
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 TGCCGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTG 1260

Qy     1261 GACTGCACCGAGCGCCAGGCGCGCCAGAACCGGACAGCAACTTCCTGGGCCCCAGCCACAAGGGCCGC 1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 GACTGCACCGAGCGCCAGGCGCGCCAGAACCGGACAGCAACTTCCTGGGCCCCAGCCACAAGGGCCGC 1320

Qy     1321 CCGGCCAACTTCCTGCAGAACCGGCAGCAGGCCGGCCGCCCCGCCCCCCCCCCC 1380
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 CCGGCCAACTTCCTGCAGAACCGGCAGCAGGCCGGCCGCCCCGCCCCCCCCCCC 1380

Qy     1381 CCCCCGGAGAGCTTCCGGTTCGGCGAGGAGACCACCACCCCCTCCCAAGCCCAAGGAC 1440
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 CCCCCGGAGAGCTTCCGGTTCGGCGAGGAGACCACCACCCCCTCCCAAGCCCAAGGAC 1440

Qy     1441 CGCGGAGCCCTACCGGGCCCCTGACCCGCCGGCCCTGCTCGTTCCGGCAGCCCCCCCTG 1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 CGCGGAGCCCTACCGGGCCCCTGACCCGCCGGCCCTGCTCGTTCCGGCAGCCCCCCCTG 1500

Qy     1501 AGCCAGTAA 1509
            |||||||||
Db     1501 AGCCAGTAA 1509

RESULT 5

LOCUS          AX455935                  1494 bp     DNA        linear    PAT  06-JUL-2002
DEFINITION     Sequence 51 from Patent WO0204493.
ACCESSION      AX455935
VERSION        AX455935.1  GI:21714919
KEYWORDS       .
SOURCE         synthetic construct
  ORGANISM     synthetic construct
               artificial sequences.
REFERENCE      1
  AUTHORS      zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE        Polynucleotides encoding antigenic hiv type c polypeptides,
               polypeptides and uses thereof
  JOURNAL      Patent: WO 0204493-A 51 17-JAN-2002;
               CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..1494
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="HIV Type C Gag optimized"
ORIGIN

Query Match             86.0%;  Score 1297.6;  DB 6;  Length 1494;
Best Local Similarity   92.7%;
Matches 1402;  Conservative    0;  Mismatches   89;  Indels   21;  Gaps    3;

Qy         1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGAGAAGCTCGACAAGTGGGAGAAGATCCGC 60
             ||||||||||||||||||||| ||||||||||||||||||| |||||||||||||||||||
Db         1 ATGGGGGCCCGCGCCAGCATCCTGAGCGGCGGCAAGCTGGACGAGTGGGAGAAGATCCGC 60

Qy        61 CTCCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
             |||||||||||||||||||||||||||||  || || ||||||||| ||||||||||||||
Db        61 CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120

Qy       121 CTGGAGCGGTTCGCCGTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
             ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       121 CTGGAGCGGTTCGCCGTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180

Qy       181 ATGAAGCAGCTGCAGCCCGCCCTCCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC 240
             | ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       181 ATCAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTTCAAC 240

Qy       241 ACCGTGGCCACACCCTGTACTGCGTGCACGGCCATCGACCGACACCAAGGAGGCC 300
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       241 ACCGTGGCCACCCTGTACTGCGTGCACGCCATCGAGGTGCGCGACACCAAGGAGGCC 300

Qy       301 CTGGACAAGATCGAGGAGGAGCAGAACAAGAGTCCAGAGCAGGCCAAGGAG 360
             ||||||||||||| |||||||||||||||| ||||||||||||||||||
Db       301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGACCCAGCAGGCC 360

Qy       361 GCCGAGGCGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAATGGTG 420
             ||||||||||||||| ||||||||| ||||||||||||||||||||
Db       361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAATCTGCAGGGCCAGATGGTG 420

Qy       421 CACCAGGCCATCAGCCCTGATCGCCAGGGCCAAGGCC 480
             ||| ||||||||||||||||||||||||||||||||
Db       421 CACCAGGCCATCAGCCCCAGGACCCTGAACGCCTGGGTCAAGGTCATCGAGGAGAAGGGC 480

Qy       481 TTCAGCCCCGAGGTGATCCCCATGTTCACGGCCCTGAGCGAGGCGGACGAC 540
             ||| || ||||||||||||| ||||||||||| |||||||||||||| ||
Db       481 TTCAACCCCGAGGTGATCCCCATGTTCACGGCCCTGAGCGAGGCGGACGAC 540

RESULT 6

```
LOCUS       AX455983                1491 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 99 from Patent WO204493.
ACCESSION   AX455983
VERSION     AX455983.1  GI:21714967
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 99 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..1491
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Gag_TV2_C_ZAopt"
ORIGIN

Query Match             85.8%;  Score 1294.8;  DB 6;  Length 1491;
Best Local Similarity   92.7%;
Matches  1401;  Conservative    0;  Mismatches   87;  Indels   24;  Gaps    3;
```

(Sequence alignment — Qy = query, Db = database; the alignment consists of paired Qy/Db nucleotide lines numbered from 1..1491 on the query and 541..1491 on the database, shown in two columns. The nucleotide sequences are densely printed and not reproduced base-by-base here.)

LOCUS       AY181195                1485 bp    DNA     linear   SYN 02-OCT-2003
DEFINITION  Synthetic construct HIV-1-derived gag protein (gag) gene, complete
            cds.
ACCESSION   AY181195
VERSION     AY181195.1  GI:37413992
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1485)
  AUTHORS   Gao,F., Li,Y., Decker,J.M., Peyerl,F.W., Bibollet-Ruche,F.,
            Rodenburg,C.M., Chen,Y., Shaw,D.R., Allen,S., Musonda,R.,
            Shaw,G.M., Zajac,A.J., Letvin,N. and Hahn,B.H.
  TITLE     Codon usage optimization of HIV type 1 subtype C gag, pol, env, and
            nef genes: in vitro expression and immune responses in
            DNA-vaccinated mice.
  JOURNAL   AIDS Res. Hum. Retroviruses 19 (9), 817-823 (2003)
REFERENCE   2  (bases 1 to 1485)
  AUTHORS   Gao,F., Li,Y., Decker,J.M., Shaw,D.R., Allen,S., Musonda,R.,
            Shaw,G.M., Zajac,A.J. and Hahn,B.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2002) Medicine, Duke University Medical Center,
            112 RPIII, LaSalle Street, Durham, NC 27710, USA
FEATURES             Location/Qualifiers
     source          1..1485
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="codon usage optimization of HIV-1 96ZM651"
     gene            1..1485
                     /gene="gag"
     CDS             1..1485
                     /gene="gag"
                     /codon_start=1
                     /transl_table=11
                     /product="gag protein"
                     /protein_id="AAO65993.1"
                     /db_xref="GI:37413993"
                     /translation="MGARASILRGGKLDKWEKIRLRPGGKKRYMIKHLVWASRELERF
                     ALNPGLLETSEGCKQIIMKQLQPALQTGTEELRSLYNTVATLYCVHEGVRDTKEALD
                     RIEEEQNKIQQKTQQAKEADGKVSQNYPIVQNLQGQMVHQKLSPRTLNAWVKVIEE
                     KAFSPEVIPMFTALSEGATPQDLNTMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVH
                     AGPIAPGQMREPRGSDIAGTTSTLQEQIAWMTSNPPIPVGDIYKRWIILGLNKIVRMY
                     SPVSILDIKQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTDTLLVQNANPDCKTILK
                     ALGPGATLEEMMTACQGVGGPSHKARVLAEAMSQTNSVNILMQKSNFKGSKRMVKCFN
                     CGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQNR
                     PEPTAPPAESSFRFEETTPAPKQESKDREALTSLKSLFGSDPLSQ"
ORIGIN

Query Match      85.6%;  Score 1291.4;  DB 12;  Length 1485;
Best Local Similarity   92.6%;  Pred. No. 1.2e-127;
Matches 1403;  Conservative  0;  Mismatches 76;  Indels 36;  Gaps 3;

Qy      1  ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACAAGTGGGAGAAGATCCGC    60
           |||||||  |||  ||  ||||||||||||  |||  |||||||||  ||||||| ||||
Db      1  ATGGGGG...CGGCGGC...CAGCGCC    60

Qy     61  CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG   120
           |||||||||||||||||||  ||||| ||||||||||||||||||||||||||||  |||
Db     61  CTGCGCCCCGGCGGCAAGAAGGCCTACATGATCAAGCACCTGGTGTGGGCCAGCCGCGAG   120

Qy    121  CTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAAACCAGCGAGGGCTGCAAGCAGATC   180
           ||||||||||||||| |||  ||||||||||||||||||  ||||  ||||||||||||
Db    121  CTGGAGCGCTTGCCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC   180

Qy    181  ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC   240
           ||||||||||||||||| |||||||||||||||||||||||| ||||||||||||||||
Db    181  ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC   240

Qy    241  ACCGTGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTCCGCGACACCAAGGAGGCC   300
           ||||||||||||||||| |||||||||||  ||| ||||||| |||||||||||  ||||
Db    241  ACCGTGGCCACCCTGTACTGCGTGCACGCCGGCGTCGAGGTGCGCGACACCAAGGAGGCC   300

Qy    301  CTGGACAAGATCGAGGAGGAGCAGAACAAGATCCAGCAGAAGACCCAGCAGGCCA----   355
           |||||  |||| |||||||||||||||||||||||||  |||||||| |||||||
Db    301  CTGGACCGTGGCCACCAGAAGCTGAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATC   355

Qy    356  -AGGAGGCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAG   414
              ||  ||||||||||||||||||||| ||||||||| |||||||| ||||||||||||
Db    361  CAGGGCCCCGGCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGC   420

Qy    415  ATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAG   474
           |||||||| |||||| ||||||  |||||||||||  |||||||||||||||| |||||
Db    421  ATGGTGCACCAGAAGCTGAGCCCCCGCACCCTGGGTGAAGGTGATCGAGGAGAAGGCC    480

Qy    475  AAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCC   534
           |||||||||||||||||||| |||||||||| |||||||||  |||||||  ||||||||
Db    481  AAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCC   540

RESULT 7

Qy    781  CGGTGGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTG   840
           | ||| |||||||||||||||||||||||  ||  |||| ||||||  |||||||||||||
Db    781  CGCTGGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTG   840

Qy    841  GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC   900
           ||||| |||||||||||||||||||||||||||||| ||||||| ||||||||||||||
Db    841  GACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC   900

Qy    901  CTGCGCGCCGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTGGTG   960
           ||| |||||||||| |||||||||||||||||||||||||||||||| ||| || ||||
Db    901  CTGCGCGCCGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTGGTG   960

Qy    961  CAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGGGCCCCGGGGCCACCCTG   1020
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    961  CAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGGGCCCCGGCGCCACCCTG   1020

Qy   1021  GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTGCTG   1080
           |||||||||||||||| |||||||||||||||||||||||||| |||||||||||||||
Db   1021  GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTGCTG   1080

Qy   1081  GCCGAGGCCATGAGCCAGACCAACAGCGTGAACATCCTGATGCAGAAGAGCAACTTCAAG   1140
           |||||||||||||||||||||  |||||||||  ||||||||||||||||||||||||||
Db   1081  GCCGAGGCCATGAGCCAGACCAACAGCGTGAACATCCTGATGCAGAAGAGCAACTTCAAG   1140

Qy   1141  GGCTCCAAGCGCATGGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC   1200
           ||| |||||||||||||||||||||||||| |||||||||||| ||||| ||||||| ||
Db   1141  GGCTCCAAGCGCATGGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC   1200

Qy   1201  TGCCGCGCCCCGCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG   1260
           |||||||||| ||||||||||||||||||||||||| |||||||||||||||||||||||
Db   1201  TGCCGCGCCCCGCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG   1260

Qy   1261  GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC   1320
           |||||||||||| ||||||||||||||||||||||||||||| ||||||| ||||||||
Db   1261  GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC   1320

Qy   1321  CCCGGCAACTTCCTGCAGAACCGCCCCGAGCCCACCGCCCCGCC---GTGCCCACCGCC   1377
           ||||||||||||||||||||||||||||||||  ||||||||||   ||||||||||||
Db   1321  CCCGGCAACTTCCTGCAGAACCGCCCCGAGCCCACCGCCCCGCCGGCCGAGAGCAGCTTCCGC   1380

Qy   1378  CCCCCCGCCCCGGGCCCGGCCGTGAAGTTCAAGTTCGAGGAGACCACCCCCGCCCCCAAG   1437
           ||||||| ||||||||  ||                         ||||||||||||||
Db   1381  --------------------------CCCCCAAGCAGGAGGCCAAG   1431

Qy   1438  GACCGCGAGGCCCTACGCAGCCTGTTCGGCAGCGACCCCCTGAGCCAGTAA   1497
           |||| |        ||  ||||||||||| |||||||||||||||| ||||
Db   1432  G-------------------ACCGCGAGGCCCTGACCAGCCTGAAGTCCCTG   1479

Qy   1498  CTGAGCCAGTAA   1509
           ||||||||| ||
Db   1480  CTGAGCCAGTAA   1491

```
RESULT 8
LOCUS       AX455887               1479 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 3 from Patent WO0204493.
ACCESSION   AX455887
VERSION     AX455887.1  GI:21714880
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 3 17-JAN-2002;
            CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..1479
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="synthetic Gag of HIV strain AF110965"
ORIGIN

  Query Match     84.6%;  Score 1276.8;  DB 6;  Length 1479;
  Best Local Similarity   92.5%;
  Matches 1399;  Conservative    0;  Mismatches   77;  Indels   36;  Gaps    4;
```

```
Qy      1 ATGGGCGCCCGGCCGCCAGCATCCTTCGCGCGGCCGGAGAAGATCCGC   60
          |||| ||| ||||||||||||||| ||||||||||||||||| ||||||
Db      1 ATGGGCGCCCGGCCGCCAGCATCCTGCGCGGCCGGAGCGCCATCCGC   60
```

... (alignment continues through the sequence)

```
LOCUS       AX468543    1479 bp    DNA    linear    PAT 16-JUL-2002
DEFINITION  Sequence 63 from Patent WO0226209.
ACCESSION   AX468543
VERSION     AX468543.1  GI:21901373
KEYWORDS    .
SOURCE      Human immunodeficiency virus 1
  ORGANISM  Human immunodeficiency virus 1 (HIV-1)
            Viruses; Retro-transcribing viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1
  AUTHORS   O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
            Singh,M., Ulmer,J. and Dubensky,T.W.
  TITLE     Microparticles for delivery of the heterologous nucleic acids
  JOURNAL   Patent: WO 0226209-A 63 04-APR-2002;
            CHIRON CORPORATION (US)

FEATURES             Location/Qualifiers
     source          1..1479
                     /organism="Human immunodeficiency virus 1"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11676"

ORIGIN
RESULT 9

Query Match   84.6%;  Score 1276.8;  DB 6;  Length 1479;
Best Local Similarity   92.5%;  Pred. No. 4.1e-126;
Matches 1399;  Conservative 0;  Mismatches 77;  Indels 36;  Gaps 4;
```

Qy       1 ATGGGCGCCGGCGCGCGCAGCCATCCTTCGCGCGGCGGCGGAGAAGCTGGACAAGAGGATCCGCG   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGGGCGCCGGCGCGCGCAGCCATCCTTCGCGCGGCGGCGGAGAAGCTGGACGGCGATCCGC   60

Qy      61 CTCCGCCGCCGGCGGCCAGCATCCTTCGCGCGGCTCTCGCGGCTTCGAAGACCCGGTG  120
Db      61 CTCGCGCGGCCCGGCGCCAGAGAAGCACTACATGCGGCGGCACCTGGTGTGGGCCCAGCCGCGAG  120

Qy     121 CTGGAGGCGCTTCGCCCCTTGAAACCCCGGCGGCTTGCTGGAGACCCCGAGATC  180
Db     121 CTGGAGCAGATCAGAGACGTTCGCCCCTTGAAACCCCGGCGGCTTGCTGGAGACCCCGAGATC  180

Qy     181 ATGAAGCAGCTGCAGACCGCGCCCTTGCAGAACTAACCCCACCGCACCGAGGAGCTGCGAGCCTGTTACAAC  240
Db     181 ATCCGCCAGCTGCACCCCGCCCTTGCACCGATCGTGCAGACCCTGTTCAAC  240

Qy     241 ACCGTGGGCACCCTGTACTGGTGTCGCGGCTGCACCGTGGGTCGCGAGGTACACCAAGGAGGCC  300
Db     241 ACCGGTGGCACCCTGTACTGCTGTCGGTGCACCGTGGGTGGAAGGTGATCGAGGAGAGC  300

Qy     301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGACCCCGAGGAGATG  360
Db     301 CTGGAGGAGTCGAGGAGGAGCAGAACAAGTGCCAGCACCCAGAGGAGGGC  360

Qy     361 GCCGGAC---GGCAAGGTGAGCGAGCAGAACTACCCCACGTGCAGCTACGAGATG  417
Db     361 GCCCACAAGGGAAGGTGAGCCAGGAAGGCTGACCGAGATG  420

Qy     418 GTGCACCAGGCCATCAGCGAGCCCTGGCCTGGGTGAAGGTGATCGAGGAGAAG  477
Db     421 GTGCACCAGGCCATCAGCCCCGCCGACCCTGAACCCCTGGGTAAGCTGGAGGAGAAG  480

Qy     478 GCCTTCAGCGCCCCGAGGGTGATCCCCATGTTCCCCCAGCAGCTGAAG  537
Db     481 GCCTTCAGCCCCGGTGATCCACGTCCCCTGAACCCCTGGAGCAGCTGAAG  540

Qy     538 GACCTGAACGATGTTGAACACCGTGGGCGGCGCCCACCAGGCCCATGCAGAGATGCTGAAG  597
Db     541 GACCTGAACACGATGTTGAACACCGTGGGCGGCGCCCACCAGGCCCCGCCATGCAGAGATGCTGAAG  600

Qy     598 GACACCATCAACGAGGAGGCCCGCGAGTGGGACGCCGGCCTCACCCCGTCAGGCGCCCGGCGCCACC  657
Db     601 GACACCATCAACGAGGAGGCCCGCGAGTGGGACGCCGGCCTCACCCCGTGCAGGCGCCCGGCGCACC  660

Qy     658 GTGGCCCCGGCGCCCAGATGCGGGACGCAGGCGCAGCATCGCGCCGGCGCACCACCAGCACC  717
Db     661 ATCCCCCCGGCGCCAGATGCGGGACGCAGGCGCAGCATCGCGCCACCACCAGCACC  720

Qy     718 CTGCAGGAGGCAGATCGCGCTGGATGAGACCAGCAACCCCCGCGCGGGCCGCGACATCTAC  777
Db     721 CTGCAGGAGGCAGATCGCGCTGGATGAGACCAGCAACCCCCGCGCGGGCCGCGACATC  780

Qy     778 AAGCGGGTGGATCATCCTGGGCCCTGAACAACGAGATCGTTGCCGGACTACGTGCAGCATC  837
Db     781 AAGCGGGTGGATCATCCTGGGCCCTGAACAACGAGATCGTTGCCGGACTACGTGGAGCATC  840

Qy     838 CTGGACAATCCGCCAGGGCCTCCGGGACTACGTCAGCCGCCGCTTCTTCAAG  897
Db     841 CTGGACATCAAGCAGCGGGCCCTTCCGGGACTACGTGAGCCGCCGCTTCTTCAAG  900

Qy     898 ACCCTGCGCGGGCCGGAGCGCCAGGGCCCACCCGAGGAAGATGGGCGGAGAAGATCCCG   957
Db     901 ACCTGGCGCGGGCCGGAGGCCCAGGGCCGACCCCGAGGAGAGGCGGCGATCCGC  957

RESULT 10

AX468547

```
LOCUS       AX468547        1479 bp    DNA    linear   PAT 16-JUL-2002
DEFINITION  Sequence 67 from Patent WO0226209.
ACCESSION   AX468547
VERSION     AX468547.1  GI:21901377
KEYWORDS
SOURCE      Human immunodeficiency virus 1 (HIV-1)
  ORGANISM  Human immunodeficiency virus 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE
  AUTHORS   O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
            Singh,M., Ulmer,J. and Dubensky,T.W.
  TITLE     Microparticles for delivery of the heterologous nucleic acids
  JOURNAL   Patent: WO 0226209-A 67 04-APR-2002;
            CHIRON CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1479
                     /organism="Human immunodeficiency virus 1"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11676"
ORIGIN
```

Query Match     84.2%;  Score 1271.2;  DB 6;  Length 1479;
Best Local Similarity   91.8%;  Pred. No. 1.6e-125;
Matches 1388;  Conservative 10;  Mismatches 78;  Indels 36;  Gaps 4;

```
Qy      1  ATGGGCGCGCCGCGCGCCAGCATCCTGGCGGGCGGCGGCGCCAAGCTGGAGCGCTGGGAG   60
           ||||||||||||||||||||| |||||| ||||||||| |||||||||||||||||||||
Db      1  ATGGGGCGCGCCGCGCGCCAGCATCCTGGCGGGCGGGCGCAAGCTGGAGCGCTGGGAG    60

Qy     61  CTGCGCCCCGGCCGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  GTGCAGAACGCCAACCCGACTGCAAGACCATCCTGGGCGGCCTCTCGGCCGGGCCAGC   120

Qy    121  CTGCAGGAGCGCGGCCGGGCAAGGCGCTGGCCCAGGGCCGCCCCGGGCGCTGCAGATC   180
Db    121  CTGGAGGAGTGATGAGACCCGCCCTGCCAGGGGCGTGGGCGGCCGGCAGATC        180

Qy    181  ATGAAGCAGCTGCAGGACGTGAACATCATGATGCGAGCGAGAGCAACTTC           240
Db    181  ATCCCGCCAGCTGCACCCCGCCCTGCAGACCCGGCACCCGCGAGCGGCTTCCAAC       240

Qy    241  ACCGTGGCGCCACCCTGTTACTGGCGTGCACGCGCCGGCGACCTGAGCTC           300
Db    241  ACCGTGGCCACCCTGTTACTGCGGCCCCATCGAGGTTGCCGCGACACCAAGGAGGCC     300

Qy    301  CTGGACAAGATCGAGGACGAGGACGAGCAACAGTCCCACGCAGGAGGAGGAG           360
Db    301  CTGGACAAGATCGAAGGACTCGCCGAGGAGGACCCAGGCGCACCAGTGTCGCAGATC     360

Qy    361  GCCAAGGTGAGCGAGAGCTACCCCATGGTGCAGATGGCAGATG                  417
Db    361  GCCGCAAGGTGAAGCTGCAGAACTGCAGGGCCCAGATG                       420

Qy    418  GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGAT           477
Db    421  GTGCACCAGGCCCATCAGCCCCGCACCCTGAACGCCTGGGTGAAGGTGATTCGAGGAAG   480

Qy    478  GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGG              537
Db    481  GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGG             540

Qy    538  GACCTTGAACACGATGCTGAACACCGTGGGCGGCCACCAGGCGGCCATGCAGATGCTGAAG 597
Db    541  GACCTTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCGGCCATGCAGATGCTGAAG 600

Qy    598  GACACCATCAACGAGGAGGCCGCGGAGTGGGACGCGCTGCACCCGGTGCAGGCCGCGCACC 657
Db    601  GACACCATCAACGAGGAGGCCGCGGAGTGGGACCGCGCTGCACCCGGTGCAGGCCGCCC   660

Qy    658  GTGGGCCCCGGCCAGATGCGCGACCCCCGCGGCAGCGACATCTAC                717
Db    661  ATCGCCCCCGGCCAGATGCGCGACCCCCGCGGCAGCGACATCTAC               720

Qy    718  CTGCAGGAGCGCAGATCGCCTGGATGACCGAGGCCCCGTGGGCCACCATCTAC        777
Db    721  CTGCACGAGGCAGATCGCCTGGATGACCGAGGCCCCATCCCGTGGGCGGCCACTCTAC    780

Qy    778  AAGCGGTGAGATCATCCTGGGCGGGCTGAACAAGATCGTGCGGATGTACACGCCCGTGAGCATC 837
Db    781  AAGCCGTGGGATCATCCTGGGCGGCTGAACAAGATCGTGCGGATCGTGTACACGCCCGTGAGCATC 840

Qy    838  CTGGACATTCCGCCAGGGCGGCCCTTCGCCACTACGTGCACGC                897
Db    841  CTGGACATCAAGCACAGGGCCCCCAAGGACCCTTCCCGGCACTACGTGGACCGCTTCTTCAAG 900

Qy    898  ACCCTGCGGCCGCGGAGGACGGCCCACCCAGGACCTGAAGAACGGATGATGCCGGACCCCGCTG 957
Db    901  ACCCTGCGCCGCGGAGGACGAGCAACTGGATGAGAAGCTGAAGAACTGGATGACCCGGACACCCTGCTG 960

Qy    958  GTGCAGAACGCCAACCCGACTGCAAGACCATCCTGGGCGGCCTCTCGGCCGGGCCAGC   1017
Db    961  GTGCAGAAACGCCAACCCGACTGCAAGACCATCCTGCGGCGCTCGGCCGGCCAGC     1020

Qy   1018  CTGGAGGAGGATGATGAGACCCGCCCTGCCAGGGGCGTGGGCGGCCGGCAGATC       1077
Db   1021  CTGGAGGAGGATGATGAGACCCGCCCTGCCAGGGGCGTGGGCGGCCGGCAGATC       1080

Qy   1078  CTGGCCGGAGGCGGCGATGGCGGAAGCTGGCACACCC---AGGGTGATGATGCGCAGAAGAGCAACTTC 1137
Db   1081  CTGGCCGGAGGCGGCCAATCAGCCAGGCCAACACC---AGGGTGATGATGCGCAGAAGAGCAACTTC   1137
```

(additional alignment continues)

RESULT 11

```
LOCUS       AX455904               1479 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 20 from Patent WO0204493.
ACCESSION   AX455904
VERSION     AX455904.1  GI:21714896
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 20 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..1479
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="synthetic Gag coding sequence of HIV strain
                     AP11096S"
ORIGIN

Query Match     83.6%;  Score 1260.8;  DB 6;  Length 1479;
Best Local Similarity   91.9%;
Matches 1389;  Conservative    0;  Mismatches   87;  Indels   36;  Gaps    4;
```

```
Qy      1 ATGGGCGCGCCGCCAGCATCCTGCGGGCCGGAGAAGTGGAGAAGATCCGC 50
            |||||||||||| ||||||||||||||| || |||||||||||||||||
Db      1 ATGGGGCGGGGCCCCAGCAGCATCTCGCGGGCCGGAGAGCTGGACAAGTG 50

Qy     61 CTGGCGCCCCGGCGGCCAAAGAAGCACTACACACTGCTGGACACTGGCTG 120
            |||||||||||||||| ||| |||| ||| |||| || |||||||||||
Db     61 CTGGCGCCCCGGCGGCCAAGAAGCACTACATGCTGAAGCACCTGGTGTGG 120

Qy    121 CTGGAGAGGTTGGCCCTGACCCCGGCCAGGAGGGCTGCAAGCAGATC 180
            ||||| ||| |||||| ||||| ||||||||||||||| ||||||||
Db    121 CTGGAGAGTTCGCCCTGAACCCCGGCCTGCTGGAAGTGGCCGAGGGG 180

Qy    181 ATGAAGCAGCTGCAGCCGGCCCTGCAGACCGGCACCGAGGAGCTGAGATCT 240
            |||||||                                  ||||||
Db    181 ATCCGCCAGGTGCACCCCGGCCTTCGAGACCGGCAGCGAGGAGCTGTTCAAC 240

Qy    241 ACCGTGGCCACACCTGTATCTGCGTGCGGCGGCCGCATCGAGGTCCGC 300
            ||  |||||||||| |  ||     ||||||||||||||||||| |||
Db    241 ACCGGTGGCCATCGTGASATGTCTTCAACTGCGTSAAGGTGCCACATCGGCCC 300

Qy    301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGAGGAG 360
            ||||||||||||||||||| |||||| |||||||||||||||||||
Db    301 AACTGCCGCGCCCCCCGAAGAAGGGCTGCTGGAAAGTTCGGCAAGGAGGGG 360

Qy    361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTTCAGAATCG 417
            |||||| |||||||| |||||||||  || || |||||||||||| |
Db    361 AAGGACTGCACCGAGCGGCCAGGCCAACTTCCTGGGCCAAGATCTGGCC 417

Qy    418 GTGCACCAGGCCATCAGCCCCAGGCACCCTGAAACGGCTGGGGTGAAGGTG 477
            ||||| |||||||||| |||||    ||||||||||||| |||||| ||||
Db    418 CGCCCGGGCCAACTTCCTGCAGAGCGGCCCCCGAGCCCCACCGGACCCCCC 477

Qy    478 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCG 537
            |||||||||||||| |||||||||||||||||||||||||| ||||||
Db    478 CCCCCCCGGAGAGCGTTCCGCTTCGAGGAGATCCCCGTGATCCCCGTGCAAG 537

Qy    538 GACCTGAACACGATGTTGAACACCGTGGGGGGCCACCAGGCCGCCATGCAG 597
            ||||||||||    ||||||||||||  |||| |||||||||||| |||||
Db    538 GACCTGAACAACCATGTCGAACACCGTTGGGGGGCCACCAGGCCGCCATGCAG 597

Qy    598 GACACCAAGCAGAAGGCCATCAGGGCCCCGTGCAGCGGCCGC 657
            ||||||||   ||||||||||||| |||||||||| ||||||
Db    601 GACACCAAGCACAACGCGGAGGAGGCCGCCGAGTGGGACAGGCTGCACCCC 660

Qy    658 GTGCGCCCCGGCCAGATCCGGCGACCCCCCCGGCAGCGAGCACACC 717
            ||| ||||| ||||||||||||||||||||||||||||||| || |
Db    661 ATCCGCCCCCCCGGCCAGATGCGGGAACCCCCCGGCCAGATCCGGCCC 720

Qy    718 CTGGACAGGAGGCAGATCGCGCTCGATCGGGGATGACAGCAGCCCCCCC 777
            ||||| |||||||||| ||| ||||| || |||| |||||||||||
Db    721 CTGCAGGGACGAGATCGCGGCTGGATGACAGACCAGCAACCCCCCAAGCC 780

Qy    778 AAGGGGGTGGGATCATCCTGGGGCCTGAACAAGAACATCGTGGCGGAGTAC 837
            ||||| |||||||||||| ||||| |||||||| |||| |||| |||||||
Db    781 AAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGAATGTACAGCCCC 840

Qy    838 CTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCG 897
            |||| |||| |||||| ||||||||||| |||||||||| |||||| ||
Db    841 CTGGACATCAAGCAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACC 900

Qy    898 ACCCTGCGCGCCGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGAGACCATC 957
            |||||||||| |||| |||||||||||||||||||||||||||| ||| |||
Db    901 ACGGGACCCTGCGCGCCCCCAGAGACTGAAGAACTGGATGACCGAGACC 960

Qy    958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCC 1017
            ||||||||| ||||||||||||||||||  |||||||||| ||| |||
Db    961 GTGGAGGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCC 1020

Qy   1018 CTGGCGGGAGGATGATGAGCCGCCTGCCAGGGCGTGGGCGGCCGCCGGTG 1077
            |||| ||||||  ||||||||| |||| ||| |||||||| ||||| |||
Db   1021 CTGGCGGGAGACTGATGAGCCCCTGCCAGGGCGTGGGCGGCCGCCACCAAG 1080

Qy   1078 CTGGCCGCCGGGCGGATGAGCCAGGCGAACATCATGATGCAGGCCGGTCAAAGTTC 1137
            |||||| ||||| ||||||||| ||| || |||||||||||| ||| ||||
Db   1081 CTGGCCGAGGCCATGAGCCAGGCCAACAGTGATGATGCAGCAGAACATGAACATG 1137

Qy   1138 AAGGGCCCCCGGCCGCCAAGGCCCCAGCCATCCCGC 1197
            ||||| || ||| |||||| ||||||||||||||||
Db   1138 AAGGGCCCCCGGCAAGAGGGGCCCCAAGAGGGCCCGC 1197

Qy   1198 AACTGCCGCGCCCCCGAAGAAGGGCTGCTGGAAGTTCGGCAAGGAGGGG 1257
            |||| ||||||||||||||||  || |||||| ||||||||||||||
Db   1198 AACTGCCGCGCCCCCGAAGAAGGGCTGCTGGAAGTGGCCGAGGGGATG 1257

Qy   1258 AAGGACTGCACCGAGCGGCCAGGCCAACTTCCTGGGCCAAGATC 1317
            |||||||||||||||||||| |||||||||||| |||||||||
Db   1258 AAGGACTGCACCGAGCGGCAGGCCAACTTCCTGGGCCAGAGATC 1317

Qy   1318 CGCCCGGGCCAACTTCCTGCAGAGCGGCCCCCGAGCCCCACCGGACCCCCC 1377
            || |||||||||| |||| |||||| |||||||  |||||||||||||||||
Db   1318 CGCCCGGGCCAACATCCCTGAGAGCGGCCCCCGAGCCCCACCGGACCCCCC 1377

Qy   1378 CCCCCCCGGAGAGCGTTCCGCTTCGAGGAGATCCCCGTGATCCCCGTGCAAG 1437
            ||||||||||||| |||||||||||| |||||||||| || ||| || |||
Db   1378 CCCCCCGGAGAGCTTCGCTTCGAGGAGATCCCCGTGATCAGGCAGGCGCAAG 1437

Qy   1438 GACCGCGGAGCCCTACCGGCGAGCCCCCTACCGCCCCCCGGCAGCGGCGGCCC 1497
            |   ||||| ||||| ||||||||||||||||| ||||||||||||||||
Db   1438 G-------------GCCAGCGCCCCTACCGGCCCCCGCCAGCGCCCTGAAGAGCC 1497

Qy   1498 CTGAGCCAGTAA 1509
            ||||||||||||
Db   1468 CTGAGCCAGTAA 1479
```

```
RESULT 12
AX149648
LOCUS           AX149648      4288 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION      Sequence 2 from Patent WO0136614.
ACCESSION       AX149648
VERSION         AX149648.1  GI:14348047
KEYWORDS        .
SOURCE          Human immunodeficiency virus
  ORGANISM      Human immunodeficiency virus
                Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
                lentivirus group.
REFERENCE       1  (bases 1 to 4288)
  AUTHORS       Shao,Y., Wagner,R., Wolf,H. and Graf,M.
  TITLE         The genome of the hiv-1 inter-subtype (c/b') and use thereof
  JOURNAL       Patent: WO 0136614-A 2 25-MAY-2001;
                Geneart GMBH Gessellschaft fuer angewandte Boitechnologie (DE) ;
                Shao, Yiming  (CN)
FEATURES             Location/Qualifiers
     source          1..4288
                     /organism="Human immunodeficiency virus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:12721"

ORIGIN
Query Match    79.7%;  Score 1202.2;  DB 6;  Length 4288;
Best Local Similarity   88.5%;  Pred. No. 2.3e-118;
Matches 1335; Conservative   0; Mismatches 153; Indels  21; Gaps   2;
```

Qy        1 ATGGGGCCCGGCCCAGCATCCTGCGCGGCGGGAGAAGTGGACAAAGTGGGAGGAGAGATCCGC 60
            |||||| |||||||||||| ||||||| | ||||||||| ||| ||||||||||||||||||
Db       13 ATGGGGCGGCGCCAGGCGCCGCATCCTGAGCGGCGGCAAGCTGGCAGAAGTGGGAGAGATCAGG 72

Qy       61 CTGCGCCCCGCGGGGGCCAAGAAGCACTACACTATGCTTGGTGTGTGGGCCAGCCGGCAG 120
            |||| |||||||| |||||||||||| ||||| |||||| |||||||||| |||||||| |||
Db       73 CTGAGGCCCGGCGGGCAAGAAGCACTACATCGTGAAGAGCATCGGCCAGGGAG 132

Qy      121 CTGGAGGGGCTTCGCCCCTGAACCCCGGCCGCTGTGAGGACACGCCCGGCTGTGCCAAGCACAGATC 180
            |||||| ||| |||||||||||| || |||||||||| || ||||| ||| | ||||| |||||||
Db      133 CTGGAGAGGTTCGGCCTTGAACCCCGGCCGCTGCGAGACCAGCCAGCCCTGTGAGAGCCAGATC 192

Qy      181 ATGAAGACAGCTGCAGCCCGCCCGCTGCAGACCGGCGCACCGAGGAGTGCGCAGCCTGTACAAC 240
            |||||| |||| |||||| |||||| ||||| ||||| || |||| || ||||||||||||||
Db      193 ATGAAGACAGCTGCAGCCCGCCCCTGCAGACCGGCGCACCGAGGAGTGCGCAGCCTTGTTCAAC 252

Qy      241 ACCGTGGCCACCCTGTACTGGTGCAGTCCGGACTCCGGACGCCATCGAGGTC 300
            |||||||||| ||| || |||||| ||| |||||||||| |||||||||||
Db      253 ACCGTGGCCACCCCCTACTGCGTGCACACCGAGCTCGAGAGGCAGGGCCAAGGAGGCC 312

Qy      301 CTGGACAAGATCGAGGAGGAGGAGAACAAGTCCCAGCAGCACCCAGGCCCAAGGAG 360
            |||||| |||| |||| |||| || ||||| |||| |||||||| ||||||||
Db      313 CTGGACAAGATCCAGGACAAGATCCAGGACAAGATCCCAGGACCACCAGGCAGAGC 372

Qy      361 GCCGACGCGCAAGGTGAGCGGCCAGAGAACTACCCCCATGCGTGCAGAACCTGCAGGGCCAGGATGGTG 420
            ||| ||||| ||||| || |||||| |||| |||||||| | |||| ||||| || |||| |||| |||
Db      373 GCCGGACGCGCAAGGCGGAGGAGGCACTACCCCCATGCTGCAGAACCTGGTTCGGCAGCGGACGTGGTG 432

Qy      421 CACCAGGCCATCAGCCTCCCCGAGCCACGCCACCCTGAACGCTGGGGTGAAAGGTGTGATCGAGGAGAAGGCC 480
            ||||||||||||| |||||| ||| ||||||||||| |||||| |||||| ||||||||||||||||||||
Db      433 CACCAGGCCATCAGCCCCAGGACCCAGAATGCATGGCCTGAGGTGAAGGTGGTGAAGGAGGAGAAGGCC 492

Qy      481 TTCAGCCCCGAGTTGATCCCCATGTTCACCGCCCTTGAGCGGCAGGACGCCAGGAC 540
            |||||||| |||| ||||||| |||||| || |||||| || ||||| || |||||
Db      493 TTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTTGAGCGGCAGGATGCCAGGAC 552

Qy      541 CTGAACACAGATGTTTGAAACACCGTGTGGCGCCAGCAGCAGATGCAGATTGGAGGAC 600
            ||||| || ||| |||||| || ||||| | ||||| ||||||| |||||| |||| |
Db      553 CTGAACACCATGCTGAACACGTGAGGCGCCAGCAGATCCTGAAGGAC 612

Qy      601 ACCATCAAGGAGGAGGAGGCCGAGTGGGACCGGCTGCACGCCCCGTG 660
            ||| |||||||||||| || |||||||| ||||| ||||||| ||||
Db      613 ACCATCAAGGAGGAGGAGGCGGCCGAGTGGGACAGGCTGCACGCCATC 672

Qy      661 GCCCCGCCCCAGCATCGGCACCCCCGGCGCGGCGACCCTG 720
            ||||||||||||||||||||||| ||||||||||||||||
Db      673 GCCCCGCCCCAGCATCGGCACCCCCGGCGCGGCAACCTG 732

Qy      721 CAGGAGCAGATCCGCCTGATGCAACAGCAACCCCGTGTCCCGGACATCTACAAG 780
            ||||||||||| || |||||||| |||||| ||| |||| || |||||||||||
Db      733 CAGGAGCAGATCGGCCTGATGGGCAGCGACTACGTGACGCAGGTTCTTCAAGGAC 792

Qy      781 CGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTG 840
            |||||||||||||| ||||||||||||| || |||| || ||||||||||| ||||||
Db      793 AGGTGGATCATCCTGGGTTTAAACAAGATCGTGAAGATGTACAGCCCCACCAGCATCCTG 852

Qy      841 GACATCCGCCCAGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACGCTTCTTCAAGACC 900
            |||||||| || ||||| || ||||||||| ||||| |||||||||| ||| | |||||||
Db      853 GACATCAAGCGAGGGCCCCAAGGAGCCCTTCAGGGACTACGTGGACGCAGGTTCTTCAAGGAC 912

Qy      901 CTGCGCGCCGAGCAGGCAGGCGCCACCCAGGACGCGTGAAGAACTGGATGACCGAGGTGGGTG 960
            || |||||||| || |||||||| |||||||||| || ||||| |||||| ||||||
Db      913 CTGAGGCGCGAGCAGGCCGCACCCAGGACGCGTGAAGAACTGGATGACCGAGGTGGGTG 972

Qy      961 CAGGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCTTCGGCCCCGGCCCACCCTG 1020
            |||||| ||||||||||| |||||||| |||||| |||| || |||||||| ||||| |
Db      973 CAGGAACGCCAACCCCGACTGCAAGACCATCATCTGAGGGCCTTCGGGCCCGGCCCCATC 1032

Qy     1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCGCCCCGGCCGCCACAAGGCGGGCTGCTG 1080
            ||||||||| |||||| |||||||| || |||| ||| || ||||||||||||||||| |||||
Db     1033 GAGGAGATGATGACGGCCTGCCAGGGCGTGGGCCGCCACCAGCACCAGCAAGGTGCTG 1092

Qy     1081 GCCGAGGGCGATGAGCCAGCGCCCAGCGGCGCCCCACAGCGTGAACATCATACTCCCCGCAAGG 1140
            ||||||||||| ||| |||||||||||| ||||||| ||| |||| ||||||| |||||||
Db     1093 GCCGAGGGCCATGACCCAGGCGCCCAGCTCGGCGACCAGCAGGTGAACATCCAGAACATCTACAAG 1149

Qy     1141 GGCCCCCGGCGGCGCCCAAGCTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAAC 1200
            |||||||||| || |||| ||| |||| ||||||||||||||||||| |||||| ||||||||||
Db     1150 GGCAGCAAGGACAGCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAAC 1209

Qy     1201 TGCCCGGCCCCCGCGCAAGAAGGGCGTGCGGAAGAAGTGCGGCAAGGAGGCCCACATGAAG 1260
            ||||| ||||| ||| |||||| ||| ||||||| ||||| ||||||||| ||||| ||||
Db     1210 TGCAGGGCGCCCCAGGAGAGGAGGGCGTCGTGCTGGAAGTGCGGCAAGGAGGCCCACATCGCCAAGGAG 1269

Qy     1261 GACTGCACCGAGCCGCCAGCGCCAACTTCCTGGGCCAAGATCTTGGGCGTAGTGCGGAAGATCCTGTACAAC 1320
            |||| || ||||| || ||| ||||| ||||| ||| ||||| |||||| ||||||||||||| ||||||
Db     1270 GACTGCACGGAGCCAGCCAAGATCTGGCCAAGTTCCTGGGCCAGGTGTGGAAGAGGAGGCTGCGC 1329

Qy     1321 CCCGGGCAACTTCCTGCAGACCGGGCATCGAGGTCCGGGACCACGAGCCGCC 1380
            ||||| |||||||||| |||||| || |||| || || ||||| |||||||
Db     1330 CCCGGCCAACTTCCTGCAGAACGAG---------GCCGAGCCCACGCCCCC 1371

Qy     1381 CTGGACAGAATCGGAGGAGGAAGCGAACAACACCCCCGCCGCAGAGGCAGAGGAC 1440
            ||||| ||| || ||||| ||||| |||| ||| || |||| |||| |||||||
Db     1372 CCCGGAGGAGAGCTTCAGGTTCGGAGAGGGGACCACCAAGGCAGCAGCCGCC 1431

Qy     1441 CGCGAGCGCCTTACCCGCCGGGCCCCCCCCTG 1472
            |||     ||  ||||||||||||||||| |
Db     1432 ATCGACAAGGAGACTGTACCCCCTGACCAGCCTGAAGAGCCTGTTCGGCAGCGACCCCAGC 1491

Qy     1318 CGCCCGGCAACTTCTGCAGAACCCGC 1377
            |||||    || |||  ||  |||||
Db     1318 CGCCCCGGCAACTTCCTGCAGAACCGC 1359

Qy     1318 CCCCCCGGCCGACAGCTTCCGTGACGCG------ 1359
            |||||||||| ||   |||  ||   ||
Db     1318 CCCCGGACCCACCGCG 1377

Qy     1378 CCCCCCGGAGAGCTTCCGCTTCGGTGGAGGACCAGCAGGCAGGAGCCAAG 1437
            |||||||  |||||||  ||    ||  |||||||||||||||||||||||
Db     1360 CCCCGGAGGAGAGCTTCAGGTTCGGGAGGGGCACCACCAAGGCAGCAGC 1419

Qy     1438 GACCGCGAGCGCCTACCCGCGAGCCCCTCCGCCGCCGCCTGTTGGGCCAGCGGCGCCCC 1497
            |||||||||||| ||| |||| |||||| | ||  |||||||||||||||||| |||
Db     1420 G------------ACCCGGAGGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAGCGACCCC 1467

Qy     1498 CTGAGGCCAGTAA 1509
            |||||||||||||
Db     1468 CTGAGGCCAGTAA 1479

```
Qy   1501 AGCCAGTAA 1509
          |||||||||
Db   1492 AGCCAGGAA 1500
```

```
RESULT 13
Db
AF201927

LOCUS          AF201927              1509 bp    DNA     linear   SYN 16-MAR-2000
DEFINITION     Synthetic construct gag protein gene, complete cds.
ACCESSION      AF201927
VERSION        AF201927.1  GI:7248702
KEYWORDS       .
SOURCE         synthetic construct
  ORGANISM     synthetic construct
               artificial sequences.
REFERENCE      1  (bases 1 to 1509)
  AUTHORS      zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.B.,
               Selby,M., Otten,G.R. and Barnett,S.W.
  TITLE        Increased expression and immunogenicity of sequence-modified human
               immunodeficiency virus type 1 gag gene
  JOURNAL      J. Virol. 74 (6), 2628-2635 (2000)
  MEDLINE      20148954
  PUBMED       10684277
REFERENCE      2  (bases 1 to 1509)
  AUTHORS      zur Megede,J. and Barnett,S.W.
  TITLE        Direct Submission
  JOURNAL      Submitted (04-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton,
               Emeryville, CA 94608, USA
FEATURES             Location/Qualifiers
     source          1..1509
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     CDS             1..1509
                     /note="HIV-1SF2 p55gag"
                     /codon_start=1
                     /transl_table=11
                     /product="gag protein"
                     /protein_id="AAF43628.1"
                     /db_xref="GI:7248703"
                     /translation="MGARASVLSGGELDKWEKIRLRPGGKKKYKLKHIVWASRELERF
                     AVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIDVKDTKEALE
                     KIEEEQNKSKKKAQQAAAAAGTGNSSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKV
                     VEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVH
                     PVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIV
                     RMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNANPDCKT
                     ILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNPATIMMQRGNFRNQRKTV
                     KCFNCGKEGHTARNCRAPRKKGCWRCGREGHQMKDCTERQANFLGKIWPSYKGRPGNF
                     LQSRPEPTAPPEESFRFGEEKTTPSQKQEPIDKELYPLTSLRSLFGNDPSSQ"
ORIGIN
```

```
Query Match             79.6%;  Score 1201;  DB 12;  Length 1509;
Best Local Similarity   88.5%;  Pred. No. 4.1e-118;
Matches 1351; Conservative  0;  Mismatches 140;  Indels 36;  Gaps  3;
```

```
Qy      1 ATGGGCGCCCGCGGCAAGCACCATCCTGGCGGCGGGCGAGAAGCTGGACAAGTGGGAGAAGATCCGC 60
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGGCGCCCGCGCCTGCCGGCCGCTGCTGAGCGGCGGCGAGCTGGACAAGTGGGAGAAGATCCGC 60
```

```
Qy     61 CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 120
          ||||||||||||||||  |||||||| |||||||||||||||||||||||||||||||||
Db     61 CTGCGCCCCGGCGGCAAGAAGTACAAGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 120
```

```
Qy    121 CTGGAGAGGTTCGCCGTCGAACCCCGCGCTGCTGGTGGAGGAGCTCAAGCGAGATC 180
          ||||||||||||| |||||||||| ||||||||||||||||||||||||||||
Db    121 CTGGAGAGGCTTCGCCGTGAACCCCGCCCTGCTGGAGACACCAGCGAGATCTGG 180
```

```
Qy    181 ATGAAGCAGTGCAGCGCCCTGCAGACCGGCAGCGAGGAGCTGCGCAGCCTGTACAAC 240
          |  ||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTGGGGCCCAGCTGCAGACCGGCAGCGAGGAGCTGCGCAGCCTGTACAAC 240
```

```
Qy    241 ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTCCGCAAGACACCAAGGAGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    241 ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTCCGCAAGGACCAAGGAGGCC 300
```

```
Db    301 CTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGC 360
```

```
Db    361 GCCGGCGCCGGCACCGGCAACAGCAGCCAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 420
```

```
Qy    361 GCCGGACGGCA-------------AGGTGACCGAGAACTACCCCATCGTGCAGAACCTG 405
          ||||||||||||             |||||||||||||||||||||||||||||||||||
Db    301 CTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGC 360
```

```
Qy    406 CAGGGGCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGTG 465
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGGGGCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGTG 480
```

```
Qy    466 ATCGAGGAGGAAGGCCTTCAGCCCCATGGTGATCCCATGTTCACGGCCCTGAGCGAGGG 525
          |||||||||||||||||||||||||||||||| ||||||||| |||||||||||||||
Db    481 GTGGAGGAGGAAGGCCCTTCAGCCCCATGTTCAGCGCCCTGAGCGAGGC 540
```

```
Qy    526 GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGGCACCAGGCCGCCATG 585
          ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    541 GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGGCACCAGGCCGCCATG 600
```

```
Qy    586 CAGATGCTGAAGGAACCACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTG 645
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTG 660
```

```
Qy    646 CACGGCCCGCCCCGTGGCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC 705
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CACGGCCCGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC 720
```

```
Qy    706 GCCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCGTGCCCCTG 765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTG 780
```

```
Qy    766 GGGGACATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGGATGTACAGC 825
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GGCGAGATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGGATGTACAGC 840
```

```
Qy    826 CCCGTGAGCATCCTGGACATCCGCCAGGGCCCCAAAGAGCCCTTCCGGGACTACGTGGAC 885
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CCCACCAGCATCCTGGACATCCGCCAGGGCCCCAAAGAGCCCTTCCGGGACTACGTGGAC 900
```

```
Qy    886 CGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCGGCCACCCAGGACGTGAAGAACTGATGACC 945
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    901 CGCTTCTACAAGACCCTGCGCGCCGAGCAGGCCAGCCAGGACGTGAAGAACTGGATGACC 960
```

```
Qy    946 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTGGGC 1005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTGGGC 1020
```

```
Qy   1006 CCCGGGCGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAC 1065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCCGCGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAC 1080
```

```
Qy   1066 AAGGCCCGCGTGCTGGCCGAGGCGCGATGAGCCAGG----CCAACAGCGTGACCAACATGATG 1122
          |||||||||||||||||||||||||||||||||||    ||||||||||||||||||
Db   1081 AAGGCCCGCGTGCTGGCCGAGGCCGATGAGCCAGGGCAAGTGACAACACATGATG 1140
```

```
Qy   1123 CAGAAGAGCAACTTCAAGGGCCCCGAGGGCAAGTCAAGTGCTGCTGAAGAGGAG 1182
          |||  ||||||||||||||||||||||||||||| || ||| ||||||
Db   1141 CAGCGCGGCAACTTCCGCAACCAGCGGAAGACCGTCAAGTGCTTCAACTGCGGCAAGGAG 1200
```

```
Qy   1183 GGCCACATCGCCAAGAACTGCCGCGCCCCGAAGAAGGGCTGCTGGAAGTGCGGCAAG 1242
          ||||| |||||||||||||||||||||||||||||||||||||||| ||||
Db   1201 GGCCACACCGCCAGGAACTGCCGCGCCCCGAAGAAGGGCTGCTGGCGC 1260
```

```
Qy   1243 GAGGGGCCCACCACGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG 1302
          ||||||| |||| |||||||||||||||||||||||||||||||||||||||||||
Db   1261 GAGGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG 1320
```

```
Qy   1303 CCCAGCTACAAGGGCCGGCCCGGCAACTTCCTGCAGAGCCGGCCCGAGCCCACAGCGCCC 1362
          ||||||||||||||||||||||||||||||||||||||||||||||||||------------C 1362
Db   1321 CCCAGCTACAAGGGCCGGCCCGGCAACTTCCTGCAGAGCCGG-----------C 1362
```

RESULT 14
BD263637

LOCUS       BD263637                1515 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Improved expression of HIV polypeptides and production of
            virus-like particles.
ACCESSION   BD263637
VERSION     BD263637.1  GI:33073405
KEYWORDS    JP 2002533124-A/4.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1515)
  AUTHORS   Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
            Greer,C., Selby,M. and Walker,C.
  TITLE     Improved expression of HIV polypeptides and production of
            virus-like particles
  JOURNAL   Patent: JP 2002533124-A 4 08-OCT-2002;
            CHIRON CORP
COMMENT     OS  Artificial Sequence
            PN  JP 2002533124-A/4
            PD  08-OCT-2002
            PF  30-DEC-1999 JP 200591193
            PF  31-DEC-1998 US 60/114495,01-DEC-1999 US   60/168471 PI
            PR  SUSAN BARNETT,JAN ZUR MEGEDE, INDRESH SRIVASTAVA, YING LIAN, PI
            KARIN HARTOG,
            PI  HONG LIU, CATHERINE GREER, MARK SELBY, CHRISTOPHER WALKER PC
            C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
            C12N5/10,
            PC  C12N7/00,C12N15/02,C12N15/00,C12N5/00,A61K37/02 CC
            Description of Artificial Sequence: synthetic HIV-Gag PH  Key
            Location/Qualifiers
FEATURES             Location/Qualifiers
     source          1..1515
                     /organism='Artificial Sequence'.
                     1..1515
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Query Match     79.6%;  Score 1201;  DB 6;  Length 1515;
  Best Local Similarity   88.5%;
  Matches 1351;  Conservative    0;  Mismatches  140;  Indels   36;  Gaps  3;

Qy      1 ATGGGCGCCCGGGCGGCGAGCGTCCTGCGCGGCGGGAAGAAGCACTACATGCTGAAGCACATCGTGTGGGCCAGCCGCGAG   60
          |||||||||||||||||||||||                                                           
Db      7 ATGGGGGCCCGGGCGTCCGTGCTGAGCG                                                       66

Qy     61 CTGGCGCCCCGGGGCGGGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG  120
Db     67 CTGCGGCCCCGGGGCGGGAAGCTACAAGCTGGGGCCAGCCGCGAG  126

Qy    121 CTGGAGAGGGGCCTTCGCCCGGTGAACCCCGGCCTGCTGGAGAGCTGCAAGCAGATC  180
Db    127 CTGGAGAGGGGCTTCGCCGGTGAACCCCGGCCTGCTGGAGAACCTGGG  186

Qy    181 ATGAAGGCAGCTGCAGCCCGGCGCCGCGGATCCAGCCTGCAGACCGGCAGCGAGGAGCTGCGCAGCCTGTACAAC  246
Db    187 CTGGGCCAGCTGCAGCCGCGCCGACCGGCAGCGAGGAGCTGCGCCAGCCTGTACAAC  246

Qy    241 ACCGTGGCCACCCTGTACTGCGTGCACGAGCGCATCGAGGTCCCCGGACACCAGGAGGCC  300
Db    247 ACCGTGGGCCACCCTGTACTGCGTGCACGAGCGCATCGAGGTCCCCGGACACCAGGAGGCC  306

Qy    301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAG  360
Db    307 CTGGAGAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGCAGAAGGCCCAGGCC  366

Qy    361 GCCGACGGCA------------------AGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG  405
Db    367 GCCGCGGCAACAGCAGCCAGGTTCAGCAGAACTACCCCATCGTGCAGAACCTG  426

Qy    406 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG  465
Db    427 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG  486

Qy    466 GTGGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCCATGTTCACCGGCCCTGAGCGGGC  525
Db    487 GTGGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGC  546

Qy    526 GCCACCCCCCAGGACCTGAACACCATGTTGAACACCGTGGGGCGCCCACCAGGCCCATG  585
Db    547 GCCACCCCCCAGGACCTGAACACCATGTTGAACACCGTGGGGCGCCACCAGGCCCATG  606

Qy    586 CAGATGCTGAAGGACACCATCAACGAGGAGGCGGCGGACGTGGACCGACCCCGTG  645
Db    607 CAGATGCTGAAGGAGACCATCAACGAGGAGGCGGGGTGACCCGACGTGGACCCGTG  666

Qy    646 CACGCCGGCCCCGTGGCCCCCGGCCAGATGCGCGACCCCCGCGGCAGCGACATCGCGGC  705
Db    667 CACGCCGGCCCCATCGCCCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCGGC  726

Qy    706 GCCACCACGCCACCCTGCCAGGAGCAGATCGCCTGGATGACCAGCAACCCGTGCCGTG  765
Db    727 ACCACCAGCACCCCTGCAGGAGGCAGATCGCGTGGATGACCAACAACCCCCATCCCGTG  786

Qy    766 GGCGAGATCTACAAGCGCTGGATCATCCTGGGGCTGAACAAGATCGTGCGCATGTACAGC  825
Db    787 GGCGAGATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC  846

Qy    826 CCCGTGAGCATCCTGGACATCCGCCAGGGGCCCAAGGAGCCCTTCCGCGACTACGTGGAC  885
Db    847 CCCACCAGCATCCTGGACATCCGCCAGGGGCCCAAGGAGCCCTTCCGCGACTACGTGGAC  906

Qy    886 CGCTTCTTCAAGACCCTGCGCGCGGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACC  945
Db    907 CGCTTCTACAAGACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACC  966

Qy    946 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCCCTCGGC  1005
Db    967 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTCGGC  1026

Qy   1006 CCCGGCGCCACCCTGGAGGAGATGACCGCCTGCCAGGGCGTGGGCGGACCAC  1065
Db   1027 CCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCAC  1086

Qy   1066 AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAAGGCCATCATCATGATG  1122
Db   1087 AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGGTGACCAACCCATCATGATG  1146

Qy   1123 CAGAGAGACGCAAACTTCAAGGGCCGCAACACCGTCAAGTGCTTCAACTGCGGCAAGGAG  1182
Db   1147 CAGGGCCGGCAACTTCCGCAACACGGAGAGACCGTCAAGTGCTTCAACTGCGGCAAGGAG  1206

Qy   1183 GGCCACATCGCCAGAAACTGCCGGGCGCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAG  1242
Db   1207 GGCCACACGCCAGAAACTGCCGGGCGCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAG  1266

Qy   1243 GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCGAACTTCCTGGGGGCAAGATCTGG  1302
Db   1267 GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGGCAAGATCTGG  1326

Qy   1363 ACCGTGCCCACCCCACGCCCGCCCCCGCCGAGGAGCGTTCGGCTTCGGAGGAGGACCACCCCGCCCCC  1422
Db   1363 CCCGAGGCCACCACCCCCCCCCCCCCGGAGGAGGCTTCGGCGCGGAGGAGAGACCACCCCCC  1422

Qy   1423 AAGCACGGAGCCCCAAGGACCCTACCCGCAGCCCTGACCGCCGCCTGCGCAGCCTG  1482
Db   1423 AGCCAGAAGCAGGAGCCCATCGACAAGGAGCTGTACCCCCTGAACAGCCTGCGCAGCCTG  1482

Qy   1483 TTCGGCCAGCGCGGCCCCTTGAGCCCAGTAA  1509
          |||||||||||||||||||||||||||||||
Db   1483 TTCGGCAACGACCCCAGCAGCCCAGTAA  1509

```
RESULT 15
LOCUS       AR373320                1515 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 4 from patent US 6602705.
ACCESSION   AR373320
VERSION     AR373320.1  GI:40075423
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1515)
AUTHORS     Barnett,S.W., Megede,J., Greer,C. and Selby,M.
TITLE       Expression of HIV polypeptides and production of virus-like
            particles
JOURNAL     Patent: US 6602705-A 4 05-AUG-2003;
FEATURES             Location/Qualifiers
     source          1..1515
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match             79.6%;  Score 1201;  DB 6;  Length 1515;
Best Local Similarity   88.5%;  Pred. No. 4.1e-118;
Matches 1351;  Conservative    0;  Mismatches 140;  Indels  36;  Gaps  3;
```

```
Qy      1  ATGGGGCGCCCCGCGCATCCTGCGCGGAGGACGTGACGAAGTGGGAGAAAGATCCGC   60
           |||||| ||  || ||| ||||||||||||  ||||||||||| ||||||||||||||
Db      7  ATGGGGCGCCCGCGCGCCCAGCCGTGCTGACGCGGCGAGCTGGACAAGAGATCCGC   66

Qy     61  CTGGGCCCCGGCGGCGCAAGAGAGCACTTACATGTGGGGCGCGATGAGAGCCAGCCGGAG  120
           |||| |||||||||||||||||||||||||| ||||||||| |||||||||||||||||
Db     67  CTCGGCCCCGGCGGCGCAAGAGAGTACACGCATGGTGTGGGCCAGCCGCCCGGAG      126

Qy    121  CTGGAGGGGCTTCGCGCCCTGAACCCCGGCCTGCTGGAGACCGGCGAGGGGCTGCAGCAGATC  180
           |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127  CTGGAGCGGCTTCGCGGGTGAACCCGGCTGCTGCTGGACGAGGCGCTGCAGATC      186

Qy    181  ATGAAGCACTGCAGCCCGCCCTGAGACCCGCCTCGCGGAGGCTTCGCGGAGCTGCACCACAAC  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187  CTGGGCCAGTCGCAGCCCACCGCCCTGCGGAGCTGCGCGAGCTGCACGATCTGG      246

Qy    241  ACCGTGGCCCACCCTGTACTCGGTGCACGGGCGATGATCGGAGGTTCCGGACGAGGAGGCC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    247  ACCGTGGCCCACCCTGTACTCGGTGCACGGGGCATCGGAGGTCAGAGGAGGGC      306

Qy    301  CTGGACAAGATCGAGGAGAACAGAACAAGTCCACGAGACCGAGCCAGCCAAGGAG      360
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    307  CTTGGACGAGAGATCGAGGCCAGAGAAGAGCACTCCAAGAAGGCCGCGGCC      366

Qy    361  GCCCAACGGGCA-----------AGGTGAGCAGAACTACCCCATCGTTGCAGAACCTG      405
           |||||| |||||            |||||||||||||||||||||||||||
Db    367  GCCCCCGGCAACCGGCAACAGCGAACAGGTCGAGCCAGGCCATCGTTGCAGAGCCTG   426

Qy    406  CAGGGGCCAGATGGTGCACCAGGCCCATCAGCCCCTGAACGGCCTGGGTGAAGGTG      465
           |||||||||||||||||||||||||||||||||||||||||||||
Db    427  CAGGGGCCAGATGGTGCACCAGGCCCATCAGCCCCTGAAGGCCTGGGTGAAGGTG     486
```

```
Qy    466  ATGGAGGAGGAAGGCCCTTCAGCCCCATTGTTCATCGCCCCATTGATCCCAGTTC...  525
Db    487  GTGGAGGAGGAAGGCCTTCAGCCCCATTGTTCGCCCCATTGATCCCAGTTC...      546

Qy    526  GCCACCCCCCAGGACCTGAACACGGTGTTGAACACGGTGGCCGCCACCAGGCCGCCATG  585
Db    547  GCCACCCCCAGGACCTGAACACGGATGTTGAACACGGTGGCCGCCACCAGGCCGCCATG  606

Qy    586  CAGATGCTGAAGGACACCATCAACGACGAGGGGAGGCCGCCCTGCACCGTG         645
Db    607  CAGATGCTGAAGGACACCATCAACGACGAGGGGAGGCCGCCCTGCACCGTG         666

Qy    646  CACGGCCGGCCCCGTGGCCCCCGCCAGATGCGCGGCCCGGCAGCATGCCGGC       705
Db    667  CACGGCCGGCCCCATCGGCCCCGCCAGATGCGCGGCCCGGCAGCATGCCGGC       726

Qy    706  GCCACCAGCACCCCTGCCAGGAGCGAGCAGATGCGCTGGCGGAGGATGACCAACAACC   765
Db    727  ACCACCAGCACCCCTGCCAGGAGCAGATCGGCTGGGATGACCAACAACATCCGTG     786

Qy    766  GGCCACAGCATCTACAACGGCGTGGATCATCCTGGGGCCTGAACAAGATCGTGCGGA    825
Db    787  GGCCAGATCTACAACGGCGTGGATCATCCTGGGGCCTGAACAAGATCGTGATACAGC   846

Qy    826  CCGGTGAGCATCCTTGGGACAATCCCCAGGCCCCAAGGAGGCCCTTCCGCGGACTACGTGGAC  885
Db    847  CCCACCAGCATCCTTGGACATCCCCAGGGCCCCAAGGAGGCCCTTCCGCGGACTACGTGGAC  906

Qy    886  CGCTTCTTCAAGACCTTGCGGGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACC   945
Db    907  CGCTTCTACAAGACCCTGCGGGCTGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACC   966

Qy    946  GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC  1005
Db    967  GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC  1026

Qy   1006  CCCGGGGCGCCACCCTGGAGGAAGGATGATGCACCGGCTGCCAGGCCGGCGGCCCGGGCCAC  1065
Db   1027  ACCACCAGACCCCGCCACCCTGGAGGAGATGATGAACGCCTGCCAGGGCGTGGGGCCGGGC  1086

Qy   1066  AAGGCCCGCGTGCTGGCGGAGGCGGATGAGCCAGTG---CCAACACGGTGAAACATCATGATG  1122
Db   1087  AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGTGACGGAACCCGGCGACCATCATGATG  1146

Qy   1123  CAGAAGAGCAACTTCAAGGGCCCCCGGGCGGCAAGAGCCGTCAAGTGCTTCAACTGCGGCAAGG  1182
Db   1147  CACGCGCGGCGCAACTTCCGGAACCAGCGGAAGACGGTCAAGTGCTTCAACTGCGGCAAGG  1206

Qy   1183  GGGCCACATCGGCCACAAGAACTGCGGCCAAGAAGGGGCTGGATGAGAGTTCGGCAAG   1242
Db   1207  GGCCACACCGGCCACCAGGAACTGCGCCGCCCAAGAAGGGCTGGCGGCTGGATGAGAGTTCGGCAAG  1266

Qy   1243  GAGGGGCCACCAGATGAAGGACTGCACCGAGCGGCCAGCGCAGGGCCCAACTTCCTGGGCCAAGATCTGG  1302
Db   1267  GAGGGCCACCAGATGAAGGACTGCACCGAGCGGCAGGGCCCAACTTCCTGGGCCAAGATCTGG         1326

Qy   1303  CCCAGCGTACAAGGGCCGGCGGCCCGCCCCCGCCCCGCGGCCCCGCC-------------C  1362
Db   1327  CCCAGCTACAAGGGCCGGCCCGGCCCCGCGCCCGGCCCGCGCCCCGCCCCGCCCCC       1368

Qy   1363  ACCGTGCCCACCGCCCCGCCAGCCTTCGGCTTCGGGCCTTCGGGGAGGACGAGACCACCCC  1422
Db   1369  CCCGAGCCCACCGCCCCGCCAGAGCTTCGCCGCTTCGGGAGGACGAGACCACCCC        1428

Qy   1423  AAGCAGGAGCCCAAGGAAGACCGGGAGCCCTACCCGCCCCAGGCGCCTG            1482
Db   1429  AGCCAGGAGCAGGAGCGGGAGCCCATCGACAAGGAGCTGTACCCCCTGACCAGCCTG    1488

Qy   1483  TTCGGCGAGCGGCCCCCTGAGCCCAGTAA                                 1509
Db   1489  TTCGGCAACGACCCCAGCCAGCCAGTAA                                  1515
```

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on:          May 28, 2004, 11:33:50 ; Search time 758.462 Seconds
                 (without alignments)
                 8870.358 Million cell updates/sec

Title:           US-09-475-704A-3
Perfect score:   1479
Sequence:        1 atgggcgccgcgccagccat..........acgaccccctgagccagtaa 1479

Scoring table:   IDENTITY NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2960401 seqs, 2274450654 residues

Total number of hits satisfying chosen parameters:          5920802

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq2:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Match Length | DB ID | Description |
|---|---|---|---|---|---|
| 1 | 1479 | 100.0 | 1479 | 10 | US-09-967-464-63 | Sequence 63, Appli |
| 2 | 1479 | 100.0 | 1479 | 13 | US-09-899-575-3 | Sequence 3, Appli |
| 3 | 1477.4 | 99.9 | 3162 | 15 | US-10-190-435-18 | Sequence 18, Appl |
| 4 | 1477.4 | 99.9 | 3462 | 15 | US-10-190-435-16 | Sequence 16, Appl |
| 5 | 1477.4 | 99.9 | 4419 | 15 | US-10-190-435-19 | Sequence 19, Appl |
| 6 | 1477.4 | 99.9 | 4419 | 15 | US-10-190-305A-14 | Sequence 14, Appl |
| 7 | 1477.4 | 99.9 | 4483 | 15 | US-10-190-435-35 | Sequence 35, Appl |
| 8 | 1477.4 | 99.9 | 4606 | 15 | US-10-190-435-34 | Sequence 34, Appl |
| 9 | 1477.4 | 99.9 | 4606 | 15 | US-10-190-305A-83 | Sequence 83, Appl |
| 10 | 1477.4 | 99.9 | 4615 | 15 | US-10-190-435-36 | Sequence 36, Appl |
| 11 | 1477.4 | 99.9 | 4702 | 15 | US-10-190-435-38 | Sequence 38, Appl |
| 12 | 1477.4 | 99.9 | 4716 | 15 | US-10-190-435-17 | Sequence 17, Appl |
| 13 | 1475.8 | 99.8 | 2742 | 15 | US-10-190-305A-13 | Sequence 13, Appl |
| 14 | 1475.8 | 99.8 | 2742 | 15 | US-10-190-305A-15 | Sequence 15, Appl |
| 15 | 1475.8 | 99.8 | 3930 | 15 | US-10-190-435-9 | Sequence 9, Appli |
| 16 | 1475.8 | 99.8 | 3930 | 15 | US-10-190-435-10 | Sequence 10, Appl |
| 17 | 1475.8 | 99.8 | 5145 | 15 | US-10-190-435-11 | Sequence 11, Appl |
| 18 | 1475.8 | 99.8 | 5145 | 15 | US-10-190-435-12 | Sequence 12, Appl |
| 19 | 1475.8 | 99.8 | 5145 | 15 | US-10-190-305A-12 | Sequence 12, Appl |
| 20 | 1474.8 | 99.7 | 4713 | 15 | US-10-190-435-59 | Sequence 59, Appl |
| 21 | 1474.8 | 99.7 | 4713 | 15 | US-10-190-305A-83 | Sequence 83, Appl |
| 22 | 1473.8 | 99.6 | 5184 | 15 | US-10-190-305A-82 | Sequence 82, Appl |
| 23 | 1473.8 | 99.6 | 5184 | 15 | US-10-190-435-58 | Sequence 58, Appl |
| 24 | 1473.4 | 99.6 | 5184 | 10 | US-09-967-464-67 | Sequence 67, Appl |
| 25 | 1472.4 | 99.6 | 2742 | 15 | US-10-190-435-57 | Sequence 57, Appl |
| 26 | 1472.4 | 99.5 | 2742 | 15 | US-10-190-305A-81 | Sequence 81, Appl |
| 27 | 1463 | 98.9 | 3537 | 13 | US-09-899-575-20 | Sequence 20, Appl |
| 28 | 1315.2 | 88.9 | 3531 | 15 | US-10-190-435-13 | Sequence 13, Appli |
| 29 | 1315.2 | 88.9 | 3537 | 15 | US-10-190-435-14 | Sequence 14, Appl |
| 30 | 1315.2 | 88.9 | 3537 | 15 | US-10-190-435-15 | Sequence 15, Appl |
| 31 | 1288.8 | 87.1 | 1491 | 15 | US-09-899-575-99 | Sequence 99, Appl |
| 32 | 1288 | 87.1 | 1494 | 13 | US-09-899-575-51 | Sequence 51, Appl |
| 33 | 1276.8 | 86.3 | 1509 | 10 | US-09-967-464-64 | Sequence 64, Appl |
| 34 | 1275.2 | 86.2 | 1509 | 10 | US-09-967-464-68 | Sequence 68, Appl |
| 35 | 1273.6 | 86.1 | 1509 | 10 | US-09-967-464-68 | Sequence 68, Appl |
| 36 | 1260.8 | 85.2 | 1509 | 13 | US-09-899-575-21 | Sequence 21, Appl |
| 37 | 1221.2 | 82.6 | 9166 | 13 | US-10-359-120-168 | Sequence 168, App |
| 38 | 1206 | 81.5 | 9788 | 13 | US-10-355-120-174 | Sequence 174, App |
| 39 | 1197.8 | 81.0 | 1515 | 13 | US-10-387-336-4 | Sequence 4, Appl |
| 40 | 1197.8 | 81.0 | 4472 | 13 | US-10-387-336-75 | Sequence 75, Appl |
| 41 | 1197.8 | 81.0 | 4608 | 13 | US-10-387-336-76 | Sequence 76, Appl |
| 42 | 1197.8 | 81.0 | 4689 | 13 | US-10-387-336-74 | Sequence 74, Appl |
| 43 | 1197.8 | 81.0 | 4766 | 13 | US-10-387-336-73 | Sequence 73, Appl |
| 44 | 1196.2 | 80.9 | 2799 | 15 | US-10-241-009-18 | Sequence 18, Appl |
| 45 | 1196.2 | 80.9 | 2799 | 15 | US-10-190-434B-18 | Sequence 18, Appl |

ALIGNMENTS

RESULT 1
US-09-967-464-63

; Sequence 63, Application US/09967464
; Publication No. US2003013845341
; GENERAL INFORMATION:
;  APPLICANT: O'Hagan, Derek
;  APPLICANT: Otten, Gillis
;  APPLICANT: Donnelly, John J.
;  APPLICANT: Polo, John M.
;  APPLICANT: Barnett, Susan
;  APPLICANT: Singh, Manohan
;  APPLICANT: Ulmer, Jeffrey
;  APPLICANT: Dubensky, Jr., Thomas W.
;  TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
;  FILE REFERENCE: PP16269.004
;  CURRENT APPLICATION NUMBER: US/09/967,464
;  CURRENT FILING DATE: 2002-04-11
;  PRIOR APPLICATION NUMBER: 60/236,105
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: 60/315,905
;  PRIOR FILING DATE: 2001-08-30
;  NUMBER OF SEQ ID NOS: 68
;  SOFTWARE: Patentin version 3.1
;  SEQ ID NO 63
;  LENGTH: 1479
;  TYPE: DNA
;  ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-63

Query Match           100.0%;  Score 1479;  DB 10;  Length 1479;
Best Local Similarity 100.0%;  Pred. No. 6.9e-303;
Matches 1479;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy          1 ATGGGCGCCGCGCCAGCCATCCTGCGCCTGTTGGCCGCAGCTGGACCCCTGGGAGCCATCCGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCGCCGCGCCAGCCATCCTGCGCCTGTTGGCCGCAGCTGGACCCCTGGGAGCCATCCGC 60

RESULT 2
US-09-899-575-3

; Sequence 3, Application US/09899575
; Publication No. US20030239612A1
; GENERAL INFORMATION:
;  APPLICANT: Zur Megede, Jan
;  APPLICANT: Barnett, Susan W.
;  APPLICANT: Bgnelbrecht, Susan
;  APPLICANT: van Rensburg, Estrelita Janse
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP01631.102
;  CURRENT APPLICATION NUMBER: US/09/899,575
;  CURRENT FILING DATE: 2001-07-05
;  PRIOR APPLICATION NUMBER: 09/475,704
;  PRIOR FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 3
;  LENGTH: 1479
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
;  OTHER INFORMATION: of HIV strain AP110965
US-09-899-575-3

Query Match             100.0%;  Score 1479;  DB 13;  Length 1479;
Best Local Similarity   100.0%;  Pred. No. 6.9e-303;
Matches 1479;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

[sequence alignment data]

```
Db        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360

Qy        361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420

Qy        421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480

Qy        481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540

Qy        541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600

Qy        601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCC 660

Qy        661 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 720

Qy        721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780

Qy        781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840

Qy        841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900

Qy        901 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960

Qy        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGC 1020

Qy       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080

Qy       1081 CTGGCCGAGGCGATGAGCCAGGCCAACACCAGCGTGATGATGCAGAAGAGCAACTTCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTGGCCGAGGCGATGAGCCAGGCCAACACCAGCGTGATGATGCAGAAGAGCAACTTCAAG 1140

Qy       1141 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200

Qy       1201 TGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260

Qy       1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320

Qy       1321 CCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCCCCCGCCGAGAGCTTCCGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCCCCCGCCGAGAGCTTCCGC 1380

Qy       1381 TTCGAGGAGACCACCCCCGGCCAGAAGCAGGAGAGCAAGGACCGCGAGACCCTGACCAGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTCGAGGAGACCACCCCCGGCCAGAAGCAGGAGAGCAAGGACCGCGAGACCCTGACCAGC 1440

Qy       1441 CTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA 1479
              |||||||||||||||||||||||||||||||||||||||
Db       1441 CTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA 1479

RESULT 3
US-10-190-435-18
; Sequence 18, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT:  BARNETT, Susan W.
;  APPLICANT:  LIAN, Ying
;  APPLICANT:  ENGELBRECHT, Susan
;  APPLICANT:  VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION:  POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP18133.003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE:  2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 3162
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: GagRTmut_C
US-10-190-435-18

  Query Match             99.9%;  Score 1477.4;  DB 15;  Length 3162;
  Best Local Similarity   99.9%;  Pred. No. 1.5e-302;
  Matches 1478;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 66

Qy         61 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 126

Qy        121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 186

Qy        181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 246

Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 306

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 366

Qy        361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 426

Qy        421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 486

Qy        481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 4
US-10-190-435-16
; Sequence 16, Application US/1019045
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 3462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagProtinaRTmut_C
US-10-190-435-16

Query Match      99.9%;  Score 1477.4;  DB 15;  Length 3462;
Best Local Similarity   99.9%;  Pred. No. 1.5e-302;
Matches 1478;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy     1 ...
Db     7 ...
Qy    61 ...
Db    67 ...
Qy   121 ...
Db   127 ...
Qy   181 ...
Db   187 ...
Qy   241 ...
Db   247 ...
Qy   301 ...
Db   307 ...
Qy   361 ...
Db   367 ...
Qy   421 ...
Db   427 ...
Qy   481 ...
Db   487 ...
Qy   541 ...
Db   547 ...
Qy   601 ...
Db   607 ...
Qy   661 ...
Db   667 ...
Qy   721 ...
Db   727 ...

```
;
; FEATURE:
; ORGANISM: Artificial Sequence
; TYPE: DNA
; LENGTH: 4419
; SEQ ID NO 19
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT APPLICATION NUMBER: US/10/190,435
; PILE REFERENCE: PP18133.003 / 2302-18133
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
; APPLICANT: VAN RENSBURG, Estrelita J.
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: LIAN, Ying
; APPLICANT: BARNETT, Susan W.
; APPLICANT: ZUR MEGEDE, Jan
; GENERAL INFORMATION:
; APPLICATION NUMBER. US2003014324BA1
; Publication No. US2003014324BA1
; Sequence 19, Application US/10190435
US-10-190-435-19
RESULT 5
```

; OTHER INFORMATION: Description of Artificial Sequence: GagRTmutTatRevNef_C
US-10-190-435-19

Query Match 99.9%; Score 1477.4; DB 15; Length 4419;
Best Local Similarity 99.9%; Pred. No. 1.4e-302;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

The page contains two panels of nucleotide sequence alignments between Query (Qy) and Database (Db) sequences, with aligned position numbers including (left panel): 727/786, 781/840, 787/846, 841/900, 847/906, 901/960, 907/966, 961/1020, 967/1026, 1021/1080, 1027/1086, 1081/1140, 1087/1146, 1141/1200, 1147/1206, 1201/1260, 1207/1266, 1261/1320, 1267/1326, 1321/1380, 1327/1386, 1381/1440, 1387/1446, 1441/1479 (ending CTGAAGAGCCTGTTCGGCAACGAACGACCCCCTGAGCCAGTGAAA 1485).

Right panel positions: 1/60, 7/66, 61/120, 67/126, 121/180, 127/186, 181/240, 187/246, 241/300, 247/306, 301/360, 307/366, 361/420, 367/426, 421/480, 427/486, 481/540, 487/546, 541/600, 547/606, 601/660, 607/666, 661/720, 667/726, 721/780, 727/786, 781/840, 787/846, 841/900, 847/906, 901/960, 907/966, 961/1020.

RESULT 6
US-10-190-305A-14

; Sequence 14, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 4419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: GagRtmutEatRevNef_C
US-10-190-305A-14

Query Match       99.9%;   Score 1477.4;   DB 15;   Length 4419;
Best Local Similarity   99.9%;  
Matches 1478;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;



Sequence alignment (Qy = query, Db = database subject) with base positions.

RESULT 7
US-10-435-35

Sequence 35, Application US/10190435
Publication No. US20030143248A1

GENERAL INFORMATION:
  APPLICANT: ZUR MEGEDE, Jan
  APPLICANT: BARNETT, Susan W.
  APPLICANT: LIAN, Ying
  APPLICANT: ENGELBRECHT, Susan
  APPLICANT: VAN RENSBURG, Estrelita J.
  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
  FILE REFERENCE: PP18133.003 / 2302-18133
  CURRENT APPLICATION NUMBER: US/10/190,435
  CURRENT FILING DATE: 2002-12-30
  NUMBER OF SEQ ID NOS: 319
  SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 35:
  LENGTH: 4483
  TYPE: DNA
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence:
  OTHER INFORMATION: gp160mod.TV1.dV1V2-gagmod.BW965
US-10-190-435-35

Query Match             99.9%;  Score 1477.4;  DB 15;  Length 4483;
Best Local Similarity   99.9%;  Pred. No.1.4e-302;
Matches 1478;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

```
Db     4445  CTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAATAA 4483
```

RESULT 8
US-10-435-34

; Sequence 34, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: BARNETT, Susan W.
;  APPLICANT: LIAN, Ying
;  APPLICANT: ENGELBRECHT, Susan
;  APPLICANT: VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP18133-003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE: 2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 34
;  LENGTH: 4606
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:
;  OTHER INFORMATION: gp160mod.TV1.dV1-gagmod.BW965
US-10-190-435-34

Query Match             99.8%;  Score 1477.4;  DB 15;  Length 4606;
Best Local Similarity   99.8%;  Conservative  0;  Pred. No. 1.4e-302;
Matches 1478;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

```
Qy       1  ATGGGCGCCCCGGGCGCCCGCGCCGGCGGCAGCCATCGCGG... 60
            ||||||||||||||||||||||||||||||||||||||||||
Db    3128  ATGGGCGCCCCGGGCGCCCGCGCCGGCGGCAGCCATCGCGG... 3187

Qy      61  CTGCGCCCCCGGGCCAAGAAGTRCTACATGATGAAGCAACAC... 120
            ||||||||||||||||||||||||||||||||||||||||||
Db    3188  CTGCGCCCCCGGGGCAAGAAGTRCTACATGATGAAGCAACAC... 3247

Qy     121  CTGGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAAAGCAGATC 180
            ||||||||||||||||||||||||||||||||||||||||||||
Db    3248  CTGGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAAAGCAGATC 3307

Qy     181  ATCCGGCCAGCTCCACCCCGGCCTGGGTGAAGGTGATCGAGG 240
            ||||||||||||||||||||||||||||||||||||||||||
Db    3308  ATCCGGCCAGCTCGCACCCCCCTGCAGACCCGGCTGAAAGCAGATC 3367

Qy     241  ACCGTGGCCACCCTGTACTGCGTGCACGAAGATCGAGGTCCCG 300
            ||||||||||||||||||||||||||||||||||||||||||
Db    3368  ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCCG 3427

Qy     301  CTGGACAAGATCGAGGAGGAGCAGCAGAAGAGTCCAGCAGGCC 360
            ||||||||||||||||||||||||||||||||||||||||||
Db    3428  CTGGACAAGATCGAGGAGGAGCAGAACAAGAGTGCCAGCAGCC 3487

Qy     361  GCCGACAAGGGCAAGGTGAGCGAGAACTACCCATCGTCGTCA 420
            ||||||||||||||||||||||||||||||||||||||||||
Db    3488  GCCGACAAGGGCAAGGTGAGCGAGAACTACCCATCGTCGTCA 3547

Qy     421  GTGCACCAGGCCATCAGCCCCCACCCTGAAGCGTGGTGATCGAGG 480
            ||||||||||||||||||||||||||||||||||||||||||||
Db    3548  GTGCACCAGGCCATCAGCCCCCACCCTGAACGCCTGGGTGATCGAGG 3607

Qy     481  GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAG 540
            ||||||||||||||||||||||||||||||||||||||||||||
Db    3608  GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAG 3667

Qy     541  GACCTGAACACGATGTTGAACACCGTGGGCGGGCACCAGGCC 600
            ||||||||||||||||||||||||||||||||||||||||||
Db    3668  GACCTGAACACGATGTTGAACACCGTGGGCGGGCACCAGGCC 3727

Qy     601  GACACCATCAACGAGGAGGCCGCCGAGTGGGACAGGGTGCAC 660
            ||||||||||||||||||||||||||||||||||||||||||
Db    3728  GACACCATCAACGAGGAGGCCGCCGAGTGGGACAGGGTGCAC 3787

Qy     661  ATCCGCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATC 720
            ||||||||||||||||||||||||||||||||||||||||||
Db    3788  ATCCGCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATC 3847

Qy     721  CTGCAGGAGCAGCATCCTGGGCCTGAACACCGGCAGCATCTAC 780
            ||||||||||||||||||||||||||||||||||||||||||
Db    3848  CTGCAGGAGCAGCATCCTGGGCCTGAACACCGGCAGCATCTAC 3907

Qy     781  AAGCCGGTGGACATCATCCTGGGCCTGAACACAAGATGTCGCAGATC 840
            ||||||||||||||||||||||||||||||||||||||||||||
Db    3908  AAGCCGGTGGACATCATCCTGGGCCTGAACACAAGATCGCCGATC 3967

Qy     841  CTGGACATCAAGCACGAGGCCCCAAGGAGCCCTTCCGCGACTAC 900
            ||||||||||||||||||||||||||||||||||||||||||
Db    3968  CTGGACATCAAGCACGAGGCGCCCCAAGGAGCCCTTCCGCGACTAC 4027

Qy     901  ACCCTGCCGCGCGAGCGAGCAGCCAGGAGGTGAAGAACTGGATG 960
            ||||||||||||||||||||||||||||||||||||||||||
Db    4028  ACCCTGCCGCGCGAGCGAGCAGCCAGGAGGTGAAGAACTGGATG 4087

Qy     961  GTGCAGAACCCCCAACCCGACTGCAAGACCATCCTGCGCGCC 1020
            ||||||||||||||||||||||||||||||||||||||||||
Db    4088  GTGCAGAACCCCCAACCCGACTGCAAGACCATCCTGCGCGCC 4147

Qy    1021  CTGGAGCCGGAGGCAGATGAATGACCGGCCTTGCCAGGGGGT 1080
            ||||||||||||||||||||||||||||||||||||||||||
Db    4148  CTGGAGGAGATGATGACCGGCCTTGCCAGGGGGT 4207

Qy    1081  CTGGGCCGGAGCGGATGAGCCAGGCCAACACCAGCGTGAATGA 1140
            ||||||||||||||||||||||||||||||||||||||||||
Db    4208  CTGGGCCGGAGCGGATGAGCCAGGCCAACACCAGCGTGAATGA 4267

Qy    1141  GGCCCCCGGCCGCACATCGTCAAGTGCTTCAACTGCGGCAAG 1200
            ||||||||||||||||||||||||||||||||||||||||||
Db    4268  GGCCCCCGGCCGCACATCGTCAAGTGCTTCAACTGCGGCAAG 4327

Qy    1201  TGCCCGCGCCCCCCGCCGGAAGAGGCTGCTGAAGTGCGGCAAG 1260
            ||||||||||||||||||||||||||||||||||||||||||
Db    4328  TGCCCGCGCCCCCCGCAAGAAGGGCTGCTGAAGTGCGGCAAG 4387

Qy    1261  GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG 1320
            ||||||||||||||||||||||||||||||||||||||||||
Db    4388  GACTGCACCGAGCGGCCAGGCCAACTTCCTGGGCCAGATCTGG 4447

Qy    1321  CCGGCCAACTTCCTGCGGACGCGCCACCGGCCGCCACCGGCC 1380
            ||||||||||||||||||||||||||||||||||||||||||
Db    4448  CCCGGCCAACTTCCTGCGGACGCGCCACCGGCCGCCACCGGCC 4507

Qy    1381  TTCGGAGGACCCACCCCCAGAGCAGAGCAAGAGCTTCCGC 1440
            ||||||||||||||||||||||||||||||||||||||||
Db    4508  TTCGGAGGACCCACCCCCAGAGCAGAGAGCAAGAGCTTCCGC 4567

Qy    1441  CTGGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA 1479
            ||||||||||||||||||||||||||||||||||||||||
Db    4568  CTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAATAA 4606
```

RESULT 9
US-10-435-36

; Sequence 36, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: BARNETT, Susan W.
;  APPLICANT: LIAN, Ying
;  APPLICANT: ENGELBRECHT, Susan
;  APPLICANT: VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

```
;  FILE REFERENCE: PP18133.003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE: 2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: Patentin Ver. 2.0
;
SEQ ID NO 36
  LENGTH: 4615
  TYPE: DNA
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence:
  OTHER INFORMATION: gp160mod.TV1.dV2-gagmod.BW965
US-10-190-435-36
```

Query Match        99.9%;
Best Local Similarity   99.9%;  Pred. No. 1.4e-302;
Matches 1478;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

```
Qy        1 ATGGGGGGGCCCCCGGGGGCCCCAGCCATCCTGCGCGGCGGCGGCGGCGGCGGGGCCCATCCGC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3137 ATGGGGGGGCCCCCGGGGGCCCCAGCCATCCTGCGCGGCGGCGGCGGCGGCGGGGCCCATCCGC 3196
```

[Remaining sequence alignment data continues in paired Qy/Db rows from positions 61 through 4615 / 3197 through 4615, comprising dense nucleotide sequence blocks with alignment bars.]

RESULT 10
US-10-190-435-38
Sequence 38, Application US/10190435
Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: BARNETT, Susan W.
;  APPLICANT: LIAN, Ying
;  APPLICANT: ENGELBRECHT, Susan
;  APPLICANT: VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP18133.003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE: 2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: Patentin Ver. 2.0
;
SEQ ID NO 38
  LENGTH: 4702
  TYPE: DNA
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence:
  OTHER INFORMATION: gp160mod.TV1-gagmod.BW965
US-10-190-435-38
```

Query Match        99.9%;  Score 1477.4;  DB 15;  Length 4702;
```

Best Local Similarity　99.9%;　Conservative　0;　Pred. No. 1.4e-302;　Mismatches　1;　Indels　0;　Gaps　0;
Matches 1478;

```
Qy      1 ATGGGGCGGCCCGGCCGCCCAGCATCCTTCGCGCGGCGGCGGCCAAGCCTGGAGCGGCCGAG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3224 ATGGGGCGGCCCGGCCGCCCAGCATCCTTCGCGCGGCGGCGGCCAAGCCTGGAGCGGCCGAG 3283

Qy     61 CTGGCGCCCCGGGGCGGGGGCGGGCCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3284 CTGGCGCCCCGGGGCGGGGGCGGGCCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCC 3343
```

RESULT 11
US-10-190-435-17

Query Match　99.9%;　Score 1477.4;　DB 15;　Length 4716;
Best Local Similarity　99.9%;　Conservative　0;　Pred. No. 1.4e-302;　Mismatches　1;　Indels　0;　Gaps　0;
Matches 1478;

```
Qy      1 ATGGGGCGGCCCGGCCGCCCAGCATCCTTCGCGCGGCGGCGGCCAAGCCTGGAGCGGCCGC   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7 ATGGGGCGGCCCGGCCGCCCAGCATCCTTCGCGCGGCGGCGGCCAAGCCTGGAGCGGCCGC   66
```

RESULT 12
US-10-190-305A-13
; Sequence 13, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: BARNETT, Susan
;  APPLICANT: LIAN, Ying
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
;  TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: 2302-18702 / 18702.002
;  CURRENT APPLICATION NUMBER: US/10/190,305A
;  CURRENT FILING DATE: 2002-07-05
;  NUMBER OF SEQ ID NOS: 93
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
;  LENGTH: 4716
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:
;  OTHER INFORMATION: GagProteinaRTmutratRevNef_C
US-10-190-305A-13

Query Match    99.9%;  Score 1477.4;  DB 15;  Length 4716;
Best Local Similarity   99.9%;  Pred. No. 1.4e-302;
Matches 1478; Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

RESULT 13

US-10-190-435-20

; Sequence 20, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: BARNETT, Susan W.
;  APPLICANT: LIAN, Ying
;  APPLICANT: ENGELBRECHT, Susan
;  APPLICANT: VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP18133_003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE: 2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 20
;  LENGTH: 2742
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
US-10-190-435-20

Query Match             99.8%;  Score 1475.8;  DB 15;  Length 2742;
Best Local Similarity   99.8%;  Pred. No. 3.2e-302;
Matches 1477; Conservative    0; Mismatches    2;  Indels    0;  Gaps    0;

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      667 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 726

Qy      721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      727 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 786

Qy      781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      787 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 846

Qy      841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      847 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 906

Qy      901 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      907 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 966

Qy      961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGC 1026

Qy     1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1027 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1086

Qy     1081 CTGGCCGAGGCGATGAGCCAGGCCAACACCAGCGTGATGATGCAGAAGAGCAACTTCAAG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1087 CTGGCCGAGGCGATGAGCCAGGCCAACACCAGCGTGATGATGCAGAAGAGCAACTTCAAG 1146

Qy     1141 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1147 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1206

Qy     1201 TGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1207 TGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1266

Qy     1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1267 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1326

Qy     1321 CCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCCCCCGCCGAGAGCTTCCGC 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1327 CCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCCCCCGCCGAGAGCTTCCGC 1386

Qy     1381 TTCGAGGAGACCACCCCCGGCCAGAAGCAGGAGAGCAAGGACCGCGAGACCCTGACCAGC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1387 TTCGAGGAGACCACCCCCGGCCAGAAGCAGGAGAGCAAGGACCGCGAGACCCTGACCAGC 1446

Qy     1441 CTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA 1479
            |||||||||||||||||||||||||||||||||||  ||
Db     1447 CTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAAGAA 1485

RESULT 14
US-10-190-305A-15
; Sequence 15, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT:  BARNETT, Susan
;  APPLICANT:  LIAN, Ying
;  TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
;  TITLE OF INVENTION:  TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: 2302-18702 / 18702.002
;  CURRENT APPLICATION NUMBER: US/10/190,305A
;  CURRENT FILING DATE:  2002-07-05
;  NUMBER OF SEQ ID NOS: 93
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 2742
```

```
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
US-10-190-305A-15

  Query Match             99.8%;  Score 1475.8;  DB 15;  Length 2742;
  Best Local Similarity   99.9%;  Pred. No. 3.2e-302;
  Matches 1477;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 66

Qy       61 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       67 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 126

Qy      121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      127 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 186

Qy      181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      187 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 246

Qy      241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      247 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 306

Qy      301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      307 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 366

Qy      361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      367 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 426

Qy      421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      427 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 486

Qy      481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      487 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 546

Qy      541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      547 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 606

Qy      601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      607 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCC 666

Qy      661 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      667 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 726

Qy      721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      727 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 786

Qy      781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      787 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 846

Qy      841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      847 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 906

Qy      901 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      907 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 966
```

RESULT 15
US-10-190-435-9
; Sequence 9, Application US10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan W.
;   APPLICANT: LIAN, Ying
;   APPLICANT: ENGELBRECHT, Susan
;   APPLICANT: VAN RENSBURG, Estrelita J.
;   TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP18133.003 / 2302-18133
;   CURRENT APPLICATION NUMBER: US/10/190,435
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 319
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 3930
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmut_C
US-10-190-435-9

Query Match      99.8%;  Score 1475.8;  DB 15;  Length 3930;
Best Local Similarity   99.9%;  Pred. No. 3.1e-302;
Matches 1477;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

[Nucleotide sequence alignment between Query (Qy) and Database (Db) entries follows, presented in paired rows of aligned bases with position numbers.]

```
Qy    1201 TGCCGGCGGCCCCCCGCGCAAGAAGGGGCTGCTTGGAAGTGCGGGCGAAGGGGGCCACCAGGATGAAG 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1207 TGCCGGCGGCCCCCCGCAAGAAGGGGCTTGCTTGGAAGTGCGGGCGAAGGAGGGGCCACCAGGATGAAG 1266

Qy    1261 GACTGCACCGAGCGCCAGGCCAACTTCCTTGGGGCCAAGCCACGGGGCCGC 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1267 GACTGCACCGAGCGCCAGGCCAACTTCCTTGGGGCAAGATCTGGCCCAGCCACAGGGGCCGC 1326

Qy    1321 CCCGGCAACTTCCTGCGAGAGCGCGCCCCGAGCCCACCGCGCCCCCCGCGCCCGGAGCTTTCCGC 1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1327 CCCGGCAACTTCCTGCAGAGCGCGGCCCGCCCCCGAGCCCACCGCGAGCTTCCGC 1386

Qy    1381 TTCCGAGGAGGACCACCCCCCGGCCAGAAGCAGAGGAGCAAGGAGGACCGGCGAGACCCTGACCAGC 1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1387 TTCCGAGGAGGACCACCCCCCGGCCAGAAGCAGAGGAGCAAGGAGGACCGGCGAGACCCTGACCAGC 1446

Qy    1441 CTGAAGAGCCTGTTCGGGCAAACGACCCCCTGAGCCAGTAA 1479
           |||||||||||||||||||||||||||||||||||| ||
Db    1447 CTGAAGAGCCTGTTCGGGCAAACGACCCCCTGAGCCAAGAA 1485
```

Search completed: May 28, 2004, 16:09:19
Job time : 762.462 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:                May 28, 2004, 07:10:24 ; Search time 685.39 Seconds
                                               (without alignments)
                                               9167.164 Million cell updates/sec

Title:          US-09-475-704A-3
Perfect Score:  1479
Sequence:       1 atgggcgccgcgccgccagcat..........acgaccccctgagccagtaa 1479
Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched:       3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:         6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    N_Geneseq_29Jan04:*
              1:  geneseqn1980s:*
              2:  geneseqn1990s:*
              3:  geneseqn2000s:*
              4:  geneseqn2001as:*
              5:  geneseqn2001bs:*
              6:  geneseqn2002s:*
              7:  geneseqn2003as:*
              8:  geneseqn2003bs:*
              9:  geneseqn2003cs:*
             10:  geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1479 | 100.0 | 1479 | 3 | AAA51609 | Aaa51609 HIV synth |
| 2 | 1479 | 100.0 | 1479 | 5 | AAL44548 | Aal44548 HIV-1 p55 |
| 3 | 1479 | 100.0 | 1479 | 6 | ABL39953 | Abl39953 Synthetic |
| 4 | 1477.4 | 99.9 | 3162 | 9 | ADC13239 | Adc13239 DNA of HI |
| 5 | 1477.4 | 99.9 | 3462 | 9 | ADC13237 | Adc13237 DNA of HI |
| 6 | 1477.4 | 99.9 | 4419 | 9 | ACA03523 | Aca03523 Synthetic |
| 7 | 1477.4 | 99.9 | 4419 | 9 | ADC13240 | Adc13240 DNA of HI |
| 8 | 1477.4 | 99.9 | 4615 | 9 | ADC13257 | Adc13257 DNA of HI |
| 9 | 1477.4 | 99.9 | 4702 | 9 | ADC13259 | Adc13259 DNA of HI |
| 10 | 1477.4 | 99.9 | 4702 | 9 | ADC13522 | Aca03522 Synthetic |
| 11 | 1477.4 | 99.9 | 4716 | 7 | ADC13238 | Adc13238 DNA of HI |
| 12 | 1475.8 | 99.8 | 2742 | 9 | ACA03524 | Aca03524 Synthetic |
| 13 | 1475.8 | 99.8 | 2742 | 7 | ADC13241 | Adc13241 DNA of HI |
| 14 | 1475.8 | 99.8 | 3930 | 9 | ADC13230 | Adc13230 DNA of HI |
| 15 | 1475.8 | 99.8 | 3930 | 9 | ADC13231 | Adc13231 DNA of HI |
| 16 | 1475.8 | 99.8 | 3930 | 9 | ADC13232 | Adc13232 DNA of HI |
| 17 | 1475.8 | 99.8 | 5145 | 7 | ACA03521 | Aca03521 Synthetic |
| 18 | 1475.8 | 99.8 | 5145 | 9 | ADC13233 | Adc13233 DNA of HI |
| 19 | 1474.8 | 99.7 | 4713 | 7 | ACA03592 | Aca03592 Synthetic |
| 20 | 1474.8 | 99.7 | 4713 | 9 | ADC13280 | Adc13280 DNA of HI |
| 21 | 1473.8 | 99.6 | 5184 | 7 | ACA03591 | Aca03591 Synthetic |
| 22 | 1473.8 | 99.6 | 5184 | 9 | ADC13279 | Adc13279 DNA of |
| 23 | 1472.4 | 99.6 | 2742 | 7 | ACA03590 | Aca03590 Synthetic |
| 24 | 1472.4 | 99.6 | 2742 | 9 | ADC13278 | Adc13278 DNA of HI |
| 25 | 1463 | 98.9 | 1479 | 3 | AAA51625 | Aaa51625 HIV codon |
| 26 | 1463 | 98.9 | 1479 | 6 | ABL39957 | Abl39957 Synthetic |
| 27 | 1461.4 | 98.8 | 1479 | 6 | AAL44552 | Aal44552 HIV-1 p55 |
| 28 | 1424.4 | 96.3 | 4546 | 9 | ADC13255 | Adc13255 DNA of HI |
| 29 | 1347.4 | 91.1 | 4423 | 9 | ADC13256 | Adc13256 DNA of HI |
| 30 | 1315.2 | 88.9 | 3531 | 9 | ADC13234 | Adc13234 DNA of HI |
| 31 | 1315.2 | 88.9 | 3537 | 9 | ADC13236 | Adc13236 DNA of HI |
| 32 | 1315.2 | 88.9 | 3538 | 9 | ADC13235 | Adc13235 DNA of HI |
| 33 | 1288 | 87.1 | 1491 | 6 | ABL40020 | Abl40020 Synthetic |
| 34 | 1288.8 | 87.1 | 1494 | 7 | ABL39972 | Abl39972 Synthetic |
| 35 | 1276.8 | 86.3 | 1509 | 3 | AAA51610 | Aaa51610 HIV synth |
| 36 | 1276.8 | 86.3 | 1509 | 5 | AAL44549 | Aal44549 HIV-1 p55 |
| 37 | 1276.8 | 86.3 | 1509 | 6 | ABL39954 | Abl39954 Synthetic |
| 38 | 1262.4 | 85.4 | 1509 | 6 | ABL39958 | Abl39958 Synthetic |
| 39 | 1260.8 | 85.2 | 1509 | 3 | AAA51626 | Aaa51626 HIV codon |
| 40 | 1260.8 | 85.2 | 1509 | 6 | AAL44553 | Aal44553 HIV-1 p55 |
| 41 | 1221.2 | 82.6 | 9166 | 6 | ABK91616 | Abk91616 Modified |
| 42 | 1206 | 81.5 | 9788 | 6 | ABK91622 | Abk91622 Modified |
| 43 | 1205.4 | 81.5 | 4288 | 4 | AAH20868 | Aah20868 HIV-1 sub |
| 44 | 1197.8 | 81.0 | 1515 | 3 | AAA70412 | Aaa70412 Synthetic |
| 45 | 1197.8 | 81.0 | 4472 | 3 | AAA70472 | Aaa70472 HIV bicis |

ALIGNMENTS

RESULT 1
AAA51609

ID   AAA51609 standard; DNA; 1479 BP.
XX
AC   AAA51609;
XX
DT   31-OCT-2000  (first entry)
XX
DE   HIV synthetic Gag polynucleotide.
XX
KW   Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW   DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS   Human immunodeficiency virus; type-C strain AF110965.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FH
FT   CDS             1..1479
FT                   /*tag= a
FT                   /product= "Synthetic_Gag"
FT                   /note= "Codon usage pattern was modified and inhibitory
FT                   elements (INS) and RRE sites were inactivated resuling
FT                   in improved expression"
XX
PN   WO200039304-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;  99WO-US031273.
XX
PR   31-DEC-1998;  98US-0114495P.
PR   01-SEP-1999;  99US-0152195P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S,  Zur Megede J;
XX
DR   WPI; 2000-452401/39.
DR   P-PSDB; AAY96943.
XX
PT   Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env
PT   polypeptide and the polypeptide useful for immunizing a mammal especially
PT   human against HIV.
XX
PS   Claim 2; Page 92-93; 113pp; English.

```
CC   Expression cassettes comprising a polynucleotide encoding antigenic type
CC   C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful
CC   in DNA immunization, generation of packaging cell lines and production of
CC   Gag- and/or Env-containing proteins. Synthetic Env and Gag expression
CC   cassettes exhibit increased potency for induction of cytotoxic T-
CC   lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-
CC   assemble into non-infectious virus-like particles which are used as a
CC   matrix for the proper presentation of an antigen entrapped or associated
CC   to the immune system of the host
XX
SQ   Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Query Match       100.0%;   Score 1479; DB 3;  Length 1479;
Best Local Similarity   100.0%;
Matches 1479;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy      1 ATGGGCGCGCCCCAGCATCCTGCGCGGCGCGGCATGGAGCGCGGCCATCCGC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGGCGCGCCCCAGCATCCTGCGCGGCGCGGCATGGAGCGCGGCCATCCGC 60

Qy     61 CTGCGCGCCCCGGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCGCGCCCCGGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGG 120

Qy    121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCC 180

Qy    181 ATCCGCCCAGCTGCACCCCCCGCCCTGCGAGACCCGCAGCGAGGCTGTTCAAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATCCGCCCAGCTGCACCCCCCGCCCTGCGAGACCCGCAGCGAGGCTGTTCAAC 240

Qy    241 ACCGTGGCCCACCCTGTACTGCGTGCGCAGCAAGATCGAGGTCCCGGACAGATC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACCGTGGCCCACCCTGTACTGCGTGCGCAGCAAGATCGAGGTCCCGGACAGATC 300

Qy    301 CTGGAGCAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATGCAGGCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGGAGCAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATGCAGGCC 360

Qy    361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGG 420

Qy    421 GTGCCACCAGGCCATCAGCCCCCGACCCCGAGGTGATCGAGGAGAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGCCACCAGGCCATCAGCCCCCGACCCCGAGGTGATCGAGGAGAAG 480

Qy    481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCCCAG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCCCAG 540

Qy    541 GACCTGAACACAGGATTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GACCTGAACACAGGATTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600

Qy    601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACACCGTGCACCCCGTGCACGCCGGCCCC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACACCGTGCACCCCGTGCACGCCGGCCCC 660

Qy    661 ATCGCCCCCGGCCAGATGCGGGAGCCGCGGGGCAGCGACATCGCGGGCAGCACCAGCACC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ATCGCCCCCGGCCAGATGCGGGAGCCGCGGGGCAGCGACATCGCGGGCAGCACCAGCACC 720

Qy    721 CTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCGTGCCCGTGGGCGACATCTAC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCGTGCCCGTGGGCGACATCTAC 780

Qy    781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACAGCCCGACCAGCATC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACAGCCCGACCAGCATC 840

Qy    841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGTTCTTCAAG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGTTCTTCAAG 900

Qy    901 ACCCTGCGCGCCGAGCAGGCCAGCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ACCCTGCGCGCCGAGCAGGCCAGCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960

Qy    961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGCGCCAGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGCGCCAGC 1020

Qy   1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080

Qy   1081 CTGGCCGAGGCGATGAGCCAGGTGACCAACAGCGCCACCATCATGATGCAGAGAGGCAAC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CTGGCCGAGGCGATGAGCCAGGTGACCAACAGCGCCACCATCATGATGCAGAGAGGCAAC 1140

Qy   1141 GGCCCCCGGCGGCCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GGCCCCCGGCGGCCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200

Qy   1201 TGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260

Qy   1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACCAGCGCCGC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACCAGCGCCGC 1320

Qy   1321 CCCGGCAACTTCCTGCAGAGCCGCCCGGAGCCCACCGCCCCCGAGAGCCAGGAGCTTCGGC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CCCGGCAACTTCCTGCAGAGCCGCCCGGAGCCCACCGCCCCCGAGAGCCAGGAGCTTCGGC 1380

Qy   1381 TTCGGCGAGGAGACCACACCCCCGCAGAAGCAGGAGCAGAAGGACGAGGGCCTGACCAGC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TTCGGCGAGGAGACCACACCCCCGCAGAAGCAGGAGCAGAAGGACGAGGGCCTGACCAGC 1440

Qy   1441 CTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA 1479
          |||||||||||||||||||||||||||||||||||||||
Db   1441 CTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA 1479
```

```
RESULT 2
AAL44548
ID   AAL44548 standard; DNA; 1479 BP.
XX
AC   AAL44548;
XX
DT   29-AUG-2003  (revised)
DT   08-NOV-2002  (first entry)
XX
DE   HIV-1 p55gag polypeptide coding sequence 1.
XX
KW   HIV; dsi; vaccine; gene; immune response; microparticle;
KW   adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
KW   polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
KW   submicron emulsion; viral infection; bacterial infection;
KW   parasitic infection; HIV-1 p55gag polypeptide.
XX
OS   Human immunodeficiency virus 1.
XX
PN   WO200226209-A2.
XX
PD   04-APR-2002.
XX
PF   28-SEP-2001; 2001WO-US030540.
XX
PR   28-SEP-2000; 2000US-02361105P.
PR   30-AUG-2001; 2001US-0315905P.
XX
PA   (CHIR ) CHIRON CORP.
```

```
XX
PI   O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;
PI   Ulmer J, Dubensky TW;
XX
DR   WPI; 2002-519084/55.
XX
PT   A microparticle to which a biologically active macromolecule is adsorbed,
PT   for use as a vaccine composition to treat viral, bacterial or parasitic
PT   infections, comprises a polymer microparticle, a detergent and a
PT   submicron emulsion.
XX
XX
PS   Claim 72; Fig 1; 100pp; English.
XX
CC   The invention relates to a method of raising an immune response in a host
CC   animal. The method of the invention comprises administering a
CC   microparticle that has an adsorbent surface to which a first biologically
CC   active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC   microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC   acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC   a polyanhydride, a detergent, and submicron emulsion. The method
CC   microparticle of the invention is useful for immunising a host animal
CC   against viral, bacterial or parasitic infections. The present DNA
CC   sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to
CC   standardise OS field)
XX
SQ   Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1479;  DB 6;  Length 1479;
Best Local Similarity   100.0%;
Matches 1479;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
                                  Pred. No. 2.3e-185;

Qy          1 ATGGGGCCCCGCCCCAGCAATCCTGCGCGCGGCGGCGGCTGGACCCCTGGAGCGGCCATCCGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGGCCCCGCCCCAGCAATCCTGCGCGCGGCGGCGGCTGGACCCCTGGAGCGGCCATCCGC 60

Qy         61 CTGCGGCCCGGCGGCAAGAAGTCTACATGATGAAGCACACCTGGGTGGGGCCCCAGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCGGCCCGGCGGCAAGAAGTCTACATGATGAAGCACACCTGGGTGGGGCCCCAGG 120

Qy        121 CTGGAGAAGTTCGCCCTCGAAACCCGGCCCTGCTTGGAGACCAGCGAGGGCCAGATC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGAGAAGTTCGCCCTCGAAACCCGGCCCTGCTTGGAGACCAGCGAGGGCCAGATC 180

Qy        181 ATCCCGCCAGCTGCACCCCGCCCTGCAGAGAAGTTCCTACATGATGAAGCACACCTGGTTCAAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCCCGCCAGCTGCACCCCGCCCTGCAGAGAAGTTCCTACATGATGAAGCACACCTGGTTCAAC 240

Qy        241 ACCGTGGCCCACCTGTACTGCGTGCACGAGAAGATCGAGGTCCGACACCAAGGAGGCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCGTGGCCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGACACCAAGGAGGCC 300

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGAGTCCCAGCAGAAGATCCAGCAGCCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGAGTCCCAGCAGAAGATCCAGCAGCCC 360

Qy        361 GCCGGACAAGGCCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCCAGATG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCGGACAAGGCCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCCAGATG 420

Qy        421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480

Qy        481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540

Qy        541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600

Qy        601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCGGCCCC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCGGCCCC 660
```

```
Db        601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCGGCCCC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||
Qy        661 ATCGGCCCCCCGGCCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGACC 720

Db        661 ATCGGCCCCCCGGCCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGACC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        721 CTGCAGGAGGCAGATGCGCTGGATGACCAGCAACCCCCGTGGGCGCCATCTAC 780

Db        721 CTGCAGGAGGCAGATGCGCTGGATGACCAGCAACCCCCGTGGGCGCCATCTAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCGTGCAGCATC 840

Db        781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCGTGCAGCATC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 900

Db        841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        901 ACCCTGCGCGCCGAGCAGGCAGCACCCCAGGAGGTGAAGAACTGGATGACCGAGACACCCTGCTG 960

Db        901 ACCCTGCGCGCCGACCAGGCACCCCAGGAGGTGAAGAACTGGATGACCGAGACACCCTGCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCCTCGGCCCCGGGCCAGC 1020

Db        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCCTCGGCCCCGGGCCAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCACCAAGGCCCCGCGTG 1080

Db       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCACCAAGGCCCCGCGTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1081 CTGGCCGAGGCGATGAGCCAGGTCAACAACACCAGCATCATGATGCAGAGGAGCAACTTCAAG 1140

Db       1081 CTGGCCGAGGCGATGAGCCAGGTCAACAACACCAGCATCATGATGCAGAGGAGCAACTTCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1141 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGGAAC 1200

Db       1141 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGGAAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1201 TGCCGCGCCCCCCGGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260

Db       1201 TGCCGCGCCCCCCGGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320

Db       1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1321 CCCGGCAACTTCCTGCAGAGCCGCCCAGAGCCCACCGCCCCCCCGGAGGAGAGCTTCCGC 1380

Db       1321 CCCGGCAACTTCCTGCAGAGCCGCCCAGAGCCCACCGCCCCCCCGGAGGAGAGCTTCCGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1381 TTCGAGGAGACCACCCCCGCCCCCAAGCAGGAGCAGAAGGACCGCGAGCCCCTGACCAGC 1440

Db       1381 TTCGAGGAGACCACCCCCGCCCCCAAGCAGGAGCAGAAGGACCGCGAGCCCCTGACCAGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       1441 CTGAAGAGCCTGTTCGGGCAAACGACCCCCTGAGCCAGTAA 1479

Db       1441 CTGAAGAGCCTGTTCGGGCAAACGACCCCCTGAGCCAGTAA 1479
```

```
RESULT 3
ABL39953
ID   ABL39953 standard; DNA; 1479 BP.
XX
AC   ABL39953;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Synthetic Gag polynucleotide sequence SEQ ID NO:3.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
```

OS　Human immunodeficiency virus; type C.
OS　Synthetic.
XX
PN　WO20204493-A2.
XX
PD　17-JAN-2002.
XX
PF　05-JUL-2001; 2001WO-US021241.
PP
PR　05-JUL-2000; 2000US-0610313.
XX
PA　(CHIR ) CHIRON CORP.
PA　(UYST-) UNIV STELLENBOSCH.
XX
PI　Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg BJ;
XX
DR　WPI; 2002-154920/20.
XX
PT　New polynucleotides encoding antigenic HIV Type C polypeptides, useful in applications including DNA immunization or generation of packaging cell lines, particularly in gene therapy.
XX
PS　Example 1; Fig 1; 233pp; English.
XX
CC　The present invention describes expression cassettes comprising a
CC　polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC　type C polypeptides. The expression cassettes comprise any of the HIV
CC　type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC　(I). (I) have immunostimulant activity and can be used in gene therapy.
CC　The HIV type C polynucleotides are useful in applications including DNA
CC　immunisation, generation of packaging cell lines, and production of HIV
CC　Type C proteins. The polynucleotides are particularly useful in gene
CC　therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC　ABB06204 to ABB06215 represent sequences used in the exemplification of
CC　the present invention
XX
SQ　Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Query Match　100.0%;　Score 1479;　DB 6;　Length 1479;
Best Local Similarity　100.0%;
Matches 1479;　Conservative 0;　Mismatches 0;　Indels 0;　Gaps 0;

```
Qy     1 ATGGGGCGCGCCCGCCCAGCCATCTTCGCGCGGGGGCAAGCTTGGACAACGCCGGCCATCCGC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGGGGCGCGCCCGCCCAGCCATCTTCGCGCGGGGGCAAGCTTGGACAACGCCGGCCATCCGC 60

Qy    61 CTGCGCCCCCGGCGGCCAAGAAGTCTACATGAATGAAGCACCTGGTGTGTGGGGCCAGCCGCAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CTGCGCCCCCGGCGGCCAAGAAGTCTACATGAATGAAGCACCTGGTGTGTGGGGCCAGCCGCAG 120

Qy   121 CTGGAGAAGTTCGCCCTTGCCTGGACCAGCGAGGCGTTGCAAGCAGATC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CTGGAGAAGTTCGCCCTTGCCTGGACCAGCGAGGCGTTGCAAGCAGATC 180

Qy   181 ATCCGCCAGCTGCACCCCCGCCCTTGCAGACCCGGCTGAAGAGGCCTGTTCAAC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ATCCGCCAGCTGCACCCCCGCCCTTGCAGACCCGGCTGAAGAGGCCTGTTCAAC 240

Qy   241 ACCGTGGCCACCCTGTACTGGGTGCACAGAACAAGTGCCAGAGAGATCCGAGGCC 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ACCGTGGCCACCCTGTACTGGGTGCACAGAACAAGTGCCAGAGAGATCCGAGGCC 300

Qy   301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCC 360

Qy   361 GCCGACAAGGGCAAGGTGCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCCGACAAGGGCAAGGTGCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420

Qy   421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480

Qy   481 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCCGGCCCACCCCAG 540
         |||||||||||||||||||||||||||||||||||||||||||||
Db   481 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCCGGCCCACCCCAG 540

Qy   541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCATCAGACGCAGATGCTGAAG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCATCAGACGCAGATGCTGAAG 600

Qy   601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGCTGCACCCGGCCCCC 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGCTGCACCCGGCCCCC 660

Qy   661 ATCGGCCCCCGCCAGATGCGCGAGCCCAGGGGCAGCGACATCGCCGGCACCAGCACC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ATCGGCCCCCGCCAGATGCGCGAGCCCAGGGGCAGCGACATCGCCGGCACCAGCACC 720

Qy   721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780

Qy   781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840

Qy   841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGTTCTTCAAG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGTTCTTCAAG 900

Qy   901 ACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 ACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTG 960

Qy   961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGCCGCCAGC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGCCGCCAGC 1020

Qy  1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080

Qy  1081 CTGGCGGAGGCGGATGAGCCAGGCCAAACACCAGCGGTGATGATGCAGAGAGGCAAACTTCAAG 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 CTGGCGGAGGCGGATGAGCCAGGCCAAACACCAGCGGTGATGATGCAGAGAGGCAAACTTCAAG 1140

Qy  1141 GGCCCCCGGCGGCCAATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GGCCCCCGGCGGCCAATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200

Qy  1201 TGCCCCGCCCCAAGAGGGGCCGTGCTGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 TGCCCCGCCCCAAGAGGGGCCGTGCTGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260

Qy  1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320

Qy  1321 CCCGGCAACTTCCTGCAGAGCCGCCCAGAGCCCACCGCCCCGCCGCCAGAGCTTCGC 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CCCGGCAACTTCCTGCAGAGCCGCCCAGAGCCCACCGCCCCGCCGCCAGAGCTTCGC 1380

Qy  1381 TTCGAGGAGACCACCCCCGCCCAGAAGCAGGAGCAAGGACCCGGCGAGAGCTTCGGC 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 TTCGAGGAGACCACCCCCGCCCAGAAGCAGGAGCAAGGACCCGGCGAGAGCTTCGGC 1440

Qy  1441 CTGAAGAGAGCCTGTTCGGCAACGACCCCCTGAGCCCAGTAA 1479
         ||||||||||||||||||||||||||||||||||||||||||
Db  1441 CTGAAGAGAGCCTGTTCGGCAACGACCCCCTGAGCCCAGTAA 1479
```

```
XX
AC   ADC13239;
XX
DT   18-DEC-2003 (first entry)
XX
DB   DNA of HIV construct GagRTmut_C SEQ ID NO 18.
XX
DE   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds;
XX
OS   Human immunodeficiency virus.
XX
XX
PN   WO200304620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0031192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg BJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 15; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation. Generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3162 BP; 729 A; 1095 C; 975 G; 363 T; 0 U; 0 Other;

Query Match             99.9%;  Score 1477.4;  DB 9;  Length 3162;
Best Local Similarity   99.9%;  Pred. No. 3.3e-185;
Matches 1478; Conservative  0; Mismatches  1;  Indels  0;  Gaps  0;

Qy        1 ATGGGGCGCCCGGCCCAGCCATTCCTCGCGCGCCAAGCTGGACGAGGCCATCCGC     60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7 ATGGGGCGCCCGGCCCAGCCATTCCTCGCGCGCCAAGCTGGACGAGGCCATCCGC     66

Qy       61 CTGGCCCCCGGGGGCCAAGAGAAGTGCTACATGATGAAGCACACTGGTGTGGGGCCAGCCGGC    120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       67 CTGGCCCCCGGGGGCCAAGAGAAGTGCTACATGATGAAGCACACTGGTGTGGGGCCAGCCGGC    126

Qy      121 CTGGAGAAGTTCGCCCCTGAACCCCGGGCCTGGGACACCACGAGGGGCTGCAAGCAGATC    180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      127 CTGGAGAAGTTCGCCCCTGAACCCCGGGCCTGGGACACCACGAGGGGCTGCAAGCAGATC    186

Qy      181 ATCCGGCCACTGCACCCCGGCCCTGAGGACCCGGCAGGCGGCCTGTTCAAC    240
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      187 ATCCGGCCACTGCACCCCGGCCCTGAGGACCCGGCAGGCGGCCTGTTCAAC    246

Qy      241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTTCCGC    300
            |||||||||||||||||||||||||||||||||||||||||||||
Db      247 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTTCCGC    306

Qy      301 CTGGGACAAGGATCGGGAGGAGCAGAAGAAGTCCAGGAGGCC    360
            ||||||||||||||||||||||||||||||||||||||||||
Db      307 CTGGGACAAGGATGTGGAGGAGCAGAAGAAGTCCAGGAGGCC    366

Qy      361 GCCGACAAGGGCAAGGTGAGCGAAGAACTACCCCATCGTTGCAGAAGCTGCAGAGATG    420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      367 GCCGACAAGGGCAAGGTGAGCGAAGAACTACCCCATCGTTGCAGAAGCTGCAGAGATG    426

Qy      421 GTGCACCAGGCCATCAGCCCCCGCCACCCTGAACGGCCTGGGTGAAGGTGATCGAGGAGAAG    480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      427 GTGCACCAGGCCATCAGCCCCCGCCACCCTGAACGGCCTGGGTGAAGGTGATCGAGGAGAAG    486

Qy      481 GCCTTCAGCCCCGAGGGTGATCCCATGTTCACGGCCCTTGAGCGGCCGAGCCAG    540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      487 GCCTTCAGCCCCGAGGGTGATCCCATGTTCACGGCCCTTGAGCGGCCGAGCCAG    546

Qy      541 GACCTGAACACGGATGTTGAACACCGTGGGGGCGCCACCAGGCCCGCCATCGGCAGGATGCTGAAG    600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      547 GACCTGAACACGGATGTTGAACACCGTGGGGGCGCCACCAGGCCCGCCATCGGCAGGATGCTGAAG    606

Qy      601 GACACCATCAACGAGGAGGAGGCCCCCGACTGGGACGCGGTGCACCCGGTGCACGCGGCCC    660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      607 GACACCATCAACGAGGAGGAGGCCCCCGACTGGGACGCGGTGCACCCGGTGCACGCGGCCC    666

Qy      661 ATCGGCCCCCGGCCGAGATCGGCGGAGCCCCCGGCGGCAGCATCGCCCACCACCAGCACC    720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      667 ATCGCCCCCGGCCGAGATCGGCGGAGCCCCCGGCGGCAGCATCGCCCACCACCAGCACC    726

Qy      721 CTGGCAGGACAGATCGGCTGGATGACAGCAAGCCCCATCCCCGTGGGCGACATCTAC    780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      727 CTGGCAGGAGCAGATCGGCTGGATGACCAGCAAGCCCCATCCCCGTGGGCGACATCTAC    786

Qy      781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCGGTGGACATC    840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      787 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCGGTGAGCATC    846

Qy      841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGGTTCTTCAAG    900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      847 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGTTCTTCAAG    906

Qy      901 ACCCTGCGCGCCGAGCAGGCACCCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG    960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      907 ACCCTGCGCGCCGAGCAGGCACCCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG    966

Qy      961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCCCTGGGCCCCGCAGC    1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTTCGGGCCCCGCAGC    1026

Qy     1021 CTGGAGGAAGATGATGACCGGCCTGCGGCCTGCCAGGGCGTGGGCGGCCCACGGGCCCTG    1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1027 CTGGAGGAAGATGATGACCGGCCTGCGGCCTGCCAGGGCGTGGGCGGCCCACGGGCCGTG    1086

Qy     1081 CTGGCCGAGGGGCGGATGAGCCAGCGCAACACAACCAGGGTGATGAGTGCAGAAGAGCAACTTCAAG    1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1087 CTGGCCGAGGGGCGGATGAGCCAGCGCAACACAACCAGGGTGATGAGTGCAGAAGAGCAACTTCAAG    1146

Qy     1141 GGCCCCCCGCCGCCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGGCACATCGCCCGCAAC    1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1147 GGCCCCCCGGCCGCCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGGCACATCGCCCGCAAC    1206

Qy     1201 TGCCGGGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACAAGATGAAG    1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1207 TGCCGGGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACAAGATGAAG    1266

Qy     1261 GACTGCACCGAGCGGCCAGGCCAACTTCCTCTGGGGAAGATCTGGCCCAGGAGGGGGGCGC    1320
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1267 GACTGCACCGAGCGGCCAGGCCAACTTCCTCTGGGGAAGATCTGGCCCAGGAGGGGGGCGC    1326

Qy     1321 CCCGGGCAAACTTCCTGCAGAGCCCCCCGCCCCACCCCACCCGCCAGGAGCTTCCGC    1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1327 CCCGGGCAAACTTCCTGCAGAGCCCCCCGCCCCACCCCACCCGCCAGGAGAGGGC    1386
```

Qy   1381 TTCGGAGGAGGACCACCCCCCGGCCAGGAGGAGGACCGGGACCGGGAGGGACGGGAGGACCCTGACCAGC 1440
          |||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   1387 TTCGGAGGAGGACCACCCCCCGGCCAGGAGGAGGACCAGGAGGACCGGGAGGACCCTGACCAGC 1446

Qy   1441 CTGAAGAGCCTGTTCGGCAACCACCCCCCTGACCCAGTAA 1479
          ||||||||||||||||||||||||||||||| ||
Db   1447 CTGAAGAGCCCTGTTCGGCAACCGACCCCCCTGAGCCAGAA 1485

RESULT 5
ADC13237
ID   ADC13237 standard; DNA; 3462 BP.
XX
AC   ADC13237;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagProteinaRtmut_C SEQ ID NO 16.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO200304620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PS   Disclosure; Fig 13; 30lpp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramuscularly, intradermally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3462 BP; 790 A; 1198 C; 1068 G; 406 T; 0 U; 0 Other;

Query Match          99.9%;  Score 1477.4;  DB 9;  Length 3462;
Best Local Similarity   99.9%;  Pred. No. 3.2e-185;
Matches 1478;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy    1 ATGGGCGCCCCCCGCCCCCAGCCATCCTGCGCGGGCCCAGCCGCCGGGCCCATCCGC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7 ATGGGGCGCCCCCCGCCGCCCAGCCATCCTGCGGCCCAGCCGCCGGGCCCATCCGC 66

Qy   61 CTGCGCCCCCGGCGGCGGCCAAGAAGTGCTACATGATGAGAAGCACCTTGGTGTGGGCCAGCCGCGAG 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67 CTGCGCCCCGGCGGCGGGCAAGAAGTGCTACATGGTGCTTCAACTGCGGGCCAACATCGCCCGCAAC 126

Qy  121 CTGGAGGAAGTTCGGCTTGACCCTTGAAACCCCCGGCCTTGCTGGGAGACCAGGCAGCAGCAGATC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  127 CTGGAGGAAGTTCGGCCTGACCCTTGAAACCCCCGGCCTTGCTGGGAGACCAGGGCAGCAGCAGATC 186

Qy  181 ATCCGGCCAGCTGCACCCCGGCCCCTTGCAGACCCGGCCAGGGGCGCAGGGCAGGGCCTGTTCAAC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  187 ATCCGGCCAGCTGCACCCCGGCCCCTTGCAGACCCGGCCAGGGGCGCAGAGGGCAGGGCCTGTTCAAC 246

Qy  241 ACCGTGGCCACCCCTGTTACTGCGTTGCACGGAGAAGATCGAGGTTCCCGCACCAAGGAGGCC 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  247 ACCGTGGCCACCCCTGTTACTGCGTTGCACGGAGAAGATCGAGGTTCCCGCACCAAGGAGGCC 306

Qy  301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAAGCAGAGAAGATCCAGCAGAGGGCCGAGGCC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  307 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAAGCAGAGAAGATCCAGCAGAGGGCCGAGGCC 366

Qy  361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGGCCAGATG 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  367 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGGCCAGATG 426

Qy  421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  427 GACACCCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 486

Qy  481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACGGCCCTTGAGCCGGCCCACCCCAG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  487 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACGGCCCTTGAGCCGGCCCACCCCAG 546

Qy  541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCATGCAGATGCTGAAG 600
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  547 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCATGCAGATGCTGAAG 606

Qy  601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCGGTGCAGGCGGGCCCC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  607 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCGGTGCAGGCGGGCCCC 666

Qy  661 ATCGCCCCCGGCCAGATGCGGGAGCCCCGGGGCAGCGACATCGCCCACCAGCACC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  667 ATCGCCCCCGGCCAGATGCGGGAGCCCCGGGGCAGCGACATCGCCCACCAGCACC 726

Qy  721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  727 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 786

Qy  781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  787 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 846

Qy  841 CTGGACATCAAGCAGGGCCCAAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  847 CTGGACATCAAGCAGGGCCCAAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 906

Qy  901 ACCCTGCGCGCCGAGCAGGCCACCCCAGACGTGAAGAACTGGATGACCGACACCCTGCTG 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  907 ACCCTGCGCGCCGAGCAGGCCACCCCAGACGTGAAGAACTGGATGACCGACACCCTGCTG 966

Qy  961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCCTCGTGGGCCCCGGCCAGC 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCCTCGTGGGCCCCGGCCAGC 1026

Qy 1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1027 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1086

Qy 1081 CTGGCCGAGGCGATGAGCCAGGCCAACAGCACAGCGCTGATGATGCAGAGAGGCAACTTCAAG 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1087 CTGGCCGAGGCGATGAGCCAGGCCAACAGCACAGCGCTGATGATGCAGAGAGGCAACTTCAAG 1146

Qy 1141 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1147 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1206

```
Qy  1201  TGCCGCGGCCCCCGCGCAAGAAGGGCTGCTTGGGAAGTGCGGCAAGGAGGGGCCCACCAGATGAAG  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1207  TGCCGCGGCCCCCGCGCAAGAAGGGCTGCTTGGGAAGTGCGGCAAGGAGGGGCCCACCAGATGAAG  1266

Qy  1261  GACTGCACGAGCGGCCAGGCCAACTTCTGGCCAGTCTGGCGCCAGCCACCAAGGGGCCGC  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1267  GACTGCACGAGCGGCCAGGCCAACTTCCTGGCCAGTCTGGCGCCAGCCACCAAGGGCCGC  1326

Qy  1321  CCCGGCCAACTTCCTGCAGAGCCCCCCAGCCCTGCACCGCCCACCACCGAGAGCTTCCGC  1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1327  CCCGGCCAACTTCCTGCAGAGCCCCCCAGCCCTGCACCGCCCACCACCGAGAGCTTCCGC  1386

Qy  1381  TTCGAGGAGACCACCCCCGGCCCCGGGCCCCCCCCCAGCCCCGAGAGTGCTACATGATGAGCTTCCGC  1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1387  TTCGAGGAGACCACCCCCGGCCCCGGGCCCCCCCCCAGCCCCGAGAGTGCTACATGATGAGCTTCCGC  1446

Qy  1441  CTGAAGAGCCTGTTCGGCCAAGACACCCCCTGAGCCAGTAA  1479
          ||||||||||||||||||||||||||||||||||||||
Db  1447  CTGAAGAGCCTGTTCGGCCAAGACACCCCCTGAGCCAGAA  1485
```

```
ID   ACA03523  standard; DNA; 4419 BP.
XX
AC   ACA03523;
XX
DT   22-MAY-2003  (first entry)
XX
DE   Synthetic DNA encoding immunogenic HIV peptide #6.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
XX
PN   WO200304657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0301192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0347728P.
PR   16-JAN-2002; 2002US-0347793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Fig 11; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation of, or for generating
CC   an immune response (e.g. a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
```

RESULT 6

ACA03523

```
CC   human immunodeficiency virus immunogenic peptide
XX
SQ   Sequence 4419 BP; 980 A; 1582 C; 1371 G; 486 T; 0 U; 0 Other;

Query Match      99.9%;  Score 1477.4;  DB 7;  1371 G; 486 T;  Length 4419;
Best Local Similarity   99.9%;  Pred. No. 3.1e-185;  DB 7;  0 U; 0 Other;
Matches 1478;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;
```

```
Qy    1  ATGGGCGCCGGCGCCGGCCAGCATCCTGCGCGCGGCGGCCAAGCTGGACGGCCATCCGC   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7  ATGGGCGCCGGCGCCGGCCAGCATCCTGCGCGCGGCGGCCAAGCTGGACGGCCATCCGC   66

Qy   61  CTGCGCCCCCGGGGCGGCGCCAGAGTGCTACATGATGAGCTTCCGCCTGAAGAGCCTG   120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67  CTGCGCCCCCGGGGCGGCGCCAGAGTGCTACATGATGAGCTTCCGCCTGAAGAGCCTG   126

Qy  121  CTGGAGGAGTTCGGCCCCTGAAACCCCGGGCCCTGTCTGAGACCAGCAGCCTGTTCAAC   180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  127  CTGGAGGAGTTCGGCCCCTGAAACCCCGGGCCCTGTCTGAGACCAGCAGCCTGTTCAAC   186

Qy  181  ATCCGCCAGCTGCACCCCGGCCCCTGCAGACCGGCAGCAGGAGCAGGAGCCTGTTCAAC   240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  187  ATCCGCCAGCTGCACCCCGGCCCCTGCAGACCGGCAGCAGGAGCAGGAGCCTGTTCAAC   246

Qy  241  ACCGTGGCCCACCCTGTACTGCGTGCCACCAGAGAAGATCGAGGTCCCCGACAGCAGGCC   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  247  ACCGTGGCCCACCCTGTACTGCGTGCCACCAGAGAAGATCGAGGTCCCCGACAGCAGGCC   306

Qy  301  CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGATCCAGCACCAGAGGCC   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  307  CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGATCCAGCACCAGAGGCC   366

Qy  361  GCCCAGCAAGGGCAAGGTGAGCGAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  367  GCCCAGCAAGGGCAAGGTGAGCGAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG   426

Qy  421  GTGCACCAGGCCATCAGCCCCCGCACCCTGAACCCCGGGCCTGGGTGATCGAGGAGAAG   480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  427  GTGCACCAGGCCATCAGCCCCCGCACCCTGAACCCCGGGCCTGGGTGATCGAGGAGAAG   486

Qy  541  GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  547  GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG   606

Qy  481  GCCTTCAGCCCCGAGGTCGATCCCCATGTTCACCGGCCCCTGAGCGGCCCACCCCCCAG   540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  487  GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGGCCCCTGAGCGGCCCACCCCCCAG   546

Qy  601  GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGGTGCACCCGGCGCGCGGCCCC   660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  607  GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGGTGCACCCGGCGCGCGGCCCC   666

Qy  661  ATCCCCGTCGGGGCGGCCGAGCCCGGCCGGCGACCGGGTGCACCCCACCGCACC   720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  667  ATCCCCGTCGGGGCGGCCGAGCCCGGCCGGCGACCGGGTGCACCCCACCGCACC   726

Qy  721  CTGCGAGGACCAGATCGCCTTGGATGAACCAGCAGACCAACCCCATCCCCGTGGGCGCATCTAC   780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  727  CTGCGAGGACCAGATCGCCTTGGATGAACCAGCAGACCAACCCCATCCCCGTGGGCGCATCTAC   786

Qy  781  AAGCGGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC   840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  787  AAGCGGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC   846

Qy  841  CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG   900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  847  CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG   906

Qy  901  ACCCTGCGCGCCGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG   960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  907  ACCCTGCGCGCCGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG   966

Qy  961  GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGCCGCCAGC   1020
```

RESULT 7
ADC13240
ID   ADC13240 standard; DNA; 4419 BP.
XX
AC   ADC13240;
XX
DT   18-DEC-2003  (first entry)
XX
DE   DNA of HIV construct GagRTmutTatRevNef_C SEQ ID NO 19.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15Rnaseh; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   16-JAN-2003.
XX
PP   05-JUL-2002; 2002WO-US021420.
PD   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg BJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15Rnaseh, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX

PS   Disclosure; Fig 16; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15Rnaseh, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 4419 BP; 980 A; 1582 C; 1371 G; 486 T; 0 U; 0 Other;

Query Match     99.9%;  Score 1477.4;  DB 9;  Length 4419;
Best Local Similarity   99.9%;  Pred. No. 3.1e-185;
Matches 1478;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy      1 ATGGGGCGGGCCCGGGCCCAGCATCCTGCGGGGCGGGCGGCAAGCTGCGGGCCATCCGC 60
Db      7 ATGGGGCGGGCCCGGGCCCAGCATCCTGCGGGGCGGGCGGCAAGCTGCGGGCCATCCGC 66

Qy     61 CTGGCGCCCCGGCGGCAAGAGAGTGCTACATATGAACGACGTGGTGTGTGGGCCATCCGC 120
Db     67 CTGGCGCCCCGGCGGCAAGAGAGTGCTACATATGAACGACGTGGTGTGTGGGCCGGAG 126

Qy    121 CTGGCGAGAGTTCGGCCCTGAACCCCGCCTCGGACCCCTGCCTGCAAGCAGATC 180
Db    127 CTGGCGAGAGTTCGGCCCTGAACCCCGCCTCGGACCCCTGCCTGCAAGCAGATC 186

Qy    181 ATCCGCGCCAGCGTGCCACCCCGCGCCCGCCCTGCGAGCCGCAGCCGCTGAAGCCTGTTCAAC 240
Db    187 ATCCGCGCCAGCGTGCCACCCCGCGCCCGCCCTGCGAGCCGCAGCCGCTGAAGCCTGTTCAAC 246

Qy    241 ACCGTGGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTTCCGCGACACCAAGGAGGCC 300
Db    247 ACCGTGGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTTCCGCGACACCAAGGAGGCC 306

Qy    301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGCAGGAGGCC 360
Db    307 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGCAGGAGGCC 366

Qy    361 GCCGACAAGGGCAAGGTGAGCAGGAACTACCCCATCGTGCAGAACCTTCAGGGCCAGATG 420
Db    367 GCCGACAAGGGCAAGGTGAGCAGGAACTACCCCATCGTGCAGAACCTTCAGGGCCAGATG 426

Qy    421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
Db    427 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 486

Qy    481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCGCCCTGAGCGAGGGCGCCGCCACCCCCCAG 540
Db    487 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCGCCCTGAGCGAGGGCGCCGCCACCCCCCAG 546

Qy    541 GACCTGAACACAGCGATGTTGAACACCGTGGGCGGGCCACCGAGCGCCATGCTGAAG 600
Db    547 GACCTGAACACAGCGATGTTGAACACCGTGGGCGGGCCACCGAGCGCCATGCTGAAG 606

Qy    601 GACACCATCAACGAGGAGGCGGCCGACTGGGACATCAACCCCGTGCACGCGGGCCCCG 660
Db    607 GACACCATCAACGAGGAGGCGGCCGACTGGGACATCAACCCCGTGCACGCGGGCCCCG 666

Qy    661 ATCGGCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCGGCAGCCACCAGCACC 720
Db    667 ATCGGCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCGGCAGCCACCAGCACC 726

Qy    721 CTGCAGGAGCAGATCGCCTGGATGACAAGCCCCCCATCCCCGTGGGGGACATCTAC 780
Db    727 CTGCAGGAGCAGATCGCCTGGATGACAAGCCCCCCATCCCCGTGGGGGACATCTAC 786

Qy    781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840

Db    967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCTCTTCGGCCCGGCGGCCAGC 1026

Qy   1021 CTGGGAGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
Db   1027 CTGGGAGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1086

Qy   1081 CTGGCCGAGGCGATGAGCCAAGCCAACCAGGCCAACGCCTGATCGAGGAGCAACTTCAAG 1140
Db   1087 CTGGCCGAGGCGATGAGCCAAGCCAACCAGGCCAACGCCTGATCGAGGAGCAACTTCAAG 1146

Qy   1141 GGCCCCCGGCCAGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200
Db   1147 GGCCCCCGGCCAGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1206

Qy   1201 TGCGGCGCGCCCCCAAGAGGGGCTGCTGGAAGTGCGGCAAGGAGGGGCACCAGATGAAG 1260
Db   1207 TGCGGCGCGCCCCCAAGAGGGGCTGCTGGAAGTGCGGCAAGGAGGGGCACCAGATGAAG 1266

Qy   1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGAAGATCTGGCCCCAGCCACAAGGGCCCC 1320
Db   1267 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGAAGATCTGGCCCCAGCCACAAGGGCCCC 1326

Qy   1321 CCCGGCCAACTTCCTGCAGAGCGGCCGCCCCGGGAAGGCCCCAGCCCCGAGAGGCTTCCGC 1380
Db   1327 CCCGGCCAACTTCCTGCAGAGCGGCCGCCCCGGGAAGGCCCCAGCCCCGAGAGGCTTCCGC 1386

Qy   1381 TTCGGAGAGACCACCCCCGCCCGCCAGGAGCCCACGGGGAGAGGAGCAAGGAAGGAGTTCCCC 1440
Db   1387 TTCGGAGAGACCACCCCCGCCCGCCAGGAGCCCACGGGGAGAGGAGCAAGGAAGGAGTTCCCC 1446

Qy   1441 CTGGAAGAGGCCTGTTCGGGGCAAGCCAACCCCCTGAGCGAGTAA 1479
Db   1447 CTGGAAGAGGCCTGTTCGGGGCAAGCCAACCCCCTGAGCGAGCCAGAA 1485

```
Db   787 AAGGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 846
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   841 CTGGACATCAAGGCAGAGGCGCCCTTCCGCGCGGCGCTGGCCATGGGCGGATGTACAGCATC 900
Db   847 CTGGACATCAAGGCAGAGGCGCCCTTCCGCGCGGCGCTGGCCATGGGCGGATGTACAGCATC 906

Qy   901 ACCCTCGCGCGCGCAGCAGCAGCAGCCCCAAGGAGCGCCTGTGAAGAACTGGACGCACACCCTGCTG 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   907 ACCCTGCGCGCGCAGCAGCAGCCCCAAGGAGCGCCTGTGAAGAACTGGACGCACACCCTGCTG 966

Qy   961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCGCGGCCAGC 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCGCGGCCAGC 1026

Qy  1021 CTGGAGGATGATGACCGGCCTGCCAGGGCGGCGGCCAGCGCCAGCAACAAGCCCCGGTG 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1027 CTGGAGGATGATGACCGGCCTGCCAGGGCGGCGGCCAGCGCCAGCAACAAGCCCCGGTG 1086

Qy  1081 CTGGCCGAGGGCGATGAGCCAGGCCACACCAGCGTGATGCAGAACAGCAACTTCAAG 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1087 CTGGCCGAGGGCGATGAGCCAGGCCGTGATCCAGAAGAGCAACTTCAAG 1146

Qy  1141 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1147 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1206

Qy  1201 TGCCGGGCTCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1207 TGCCGGGCTCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1266

Qy  1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGCAGAGCCGCCCAGAGCCAGCCCACAAGGCCGC 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1267 GACTGCACCGAGCGCCAGGCCAACTTCCTGCAGAGCCGCCCAGATCTCGGCCCACAAGGGCGC 1326

Qy  1321 CTCCGCCCAGGGCGGCCCCGAGCCCACCGCCCCGCCCGAGCCGCCGAGGCGGCCTTCGGC 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1327 CCCGGCCAACTTCCTGCAGAGCCGCCCCGAGCCGACCGCCCCGCCCGAGAGCTTCCGC 1386

Qy  1381 TTCGGAGGAGGACCACCCCGGCCAGGAGCAGGAGGACGAGGAGGGCCTGCAGCAGAGATC 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1387 TTCGGAGGAGGACCACCCCGGCCAGGAGCAGGAGGACGAGGAGGGCCTGCAGCAGAGATC 1446

Qy  1441 CTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA 1479
         ||||||||||||||||||||||||||||||||||| ||
Db  1447 CTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGGAA 1485
```

RESULT 8

Query Match      99.9%;  Score 1477.4;  DB 9;  Length 4615;
Best Local Similarity   99.9%;   Pred. No. 3.1e-185;
Matches  1478;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

```
Qy     1 ATGGGCGCCCGGCGCCAGCATCCTCCGCGCGGCGCAAGCTGGACGCCATCGC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3137 ATGGGCGCCCGGCGCCAGCATCCTCCGCGCGGCGCAAGCTGGACGCCATCGC 3196

Qy    61 CTTCGCCCCGGGCGGCGCAAGAAGTGCTACATGATGAAAGCACCTGTGGGCCCGAG 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3197 CTCCGCCCCGGGCGGCGCAAGAAGTGCTACATGATGAAAGCACCTGGTGGGCCGGAG 3256

Qy   121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTTCTGGAGACCAGCGAGGGCTTGCAGCAGAGATC 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3257 CTGGAGAAGTTCGCCCTGAACCCCGGCCTTCTGGAGACCAGCGAGGGCTGCAGCAGAGATC 3316

Qy   181 ATCCGCCAGCTGCACCCCCGCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3317 ATCCGCCAGCTGCACCCCCGCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 3376

Qy   241 ACCGTGGCCACCCTGTACTGCGTGCACCAGAGAATCGAGGTCCGGACACCAAGGAGGCC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3377 ACCGTGGCCACCCTGTACTGCGTGCACCAGAGAATCGAGGTCCGGACACCAAGGAGGCC 3436

Qy   301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAAAAAACCCAGCAGGCCAAGGAGGCC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3437 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAAAAAACCCAGCAGGCCAAGGGCC 3496

Qy   361 GCCGACAAGGGCCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3497 GCCGACAAGGGCCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 3556

Qy   421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3557 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 3616

Qy   481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCGAG 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3617 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCGAG 3676

Qy   541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3677 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 3736

Qy   601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCC 660
```

```
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
DE   DNA of HIV construct gp160mod-TV1-gagmod-BW965 SEQ ID NO 38.
DT   18-DEC-2003  (first entry)
XX
AC   ADC13259;
ID   ADC13259 standard; DNA; 4702 BP.
XX

RESULT 9
ADC13259
```

```
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-03031192P.
PR   31-AUG-2001; 2001US-03168060P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a
PT   polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env,
PT   Int, Nef, p15RnaseH, Pol, Tat, Prot,
PT   or Rev polypeptide, useful for
PT   immunisation, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 35; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 4702 BP; 1057 A; 1617 C; 1371 G; 657 T; 0 U; 0 Other;

Query Match            Best Local Similarity   99.9%;  Score 1477.4; DB 9; Length 4702;
Matches 1478; Conservative    0; Mismatches    1; Indels    0; Gaps    0;
```

```
Qy          1 ATGGGCGCCCGGCCCAGCATCCTGCGCGGGCGGCTGAAGCTGGAGCGCATCGG        60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3224 ATGGGCGCCCGGCCCAGCATCCTGCGCGGGCGGCTGAAGCTGGAGCGCATCGG      3283

Qy         61 CTGGGCCCCGGCGGCGAAGAAGTTCTACATGATGAAGCACCTGGTGTGGGCCCGCGAG       120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3284 CTGGGCCCCGGCGGCGAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCCGCGAG     3343

Qy        121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTTCCTGAGACCAGCGAGGGCTCCAAGCAGATC       180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3344 CTGGAGAAGTTCGCCCTGAACCCCGGCCTTCCTGAGACCAGCGAGGGCTCCAAGCAGATC     3403

Qy        181 ATCCCGCCAGCTGCACCCGCCCTGCAGACCGGACCGGAGCTGAAGAGGCCTGTTCAAC       240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3404 ATCCCGCCAGCTGCACCCGCCCTGCAGACCGGACCGGAGCTGAAGAGGCCTGTTCAAC     3463

Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC       300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3464 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC     3523

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGCAGAAGCAGGCCGAGGCC       360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3524 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAGCAGCAGAAGCAGGCCGAGGCC     3583

Qy        361 GCCGACAAGGGCAAGGTGACCGGAGAACTACCCATGGTGCAGAACCTGCAGGGCCAGATG       420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3584 GCCGACAAGGGCAAGGTGACCGGAGAACTACCCATGGTGCAGAACCTGCAGGGCCAGATG     3643

Qy        421 GTGCACCAGGCCATGAGCCCCCGGACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG       480
```

```
Db       3737 GACACCCATCACGAGGAGGCCCC      3796
              ||||||||||||||||||||||||
Qy        661 ATGCGCCCCGGCGGCCCGGGGCGCGAGCCCGGCGGCCGGCCGAGGTGGGCGACCGCGCACC       720

Qy        661 ATGCGCCCCGGCGGCCCGGGGCGCGAGCCCGGCGGCCGGCCGAGGTGGGCGACCGCGCACC       720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3797 ATGCGCCCCGGCGGCCCGGGGCGCGAGCCCGGCGGCCGGCCGAGGTGGGCGACCGCGCACC     3856

Qy        721 CTGGCAGCAGATCGGCCTGGATGACCAGGCAACCGGCCCGGCGACCATCTAC       780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3857 CTGGCAGCAGATGCGCCTGGATGACCAGGCAACCCCATCCCCGTGGGCGACATCTAC     3916

Qy        781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC       840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3917 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC     3976

Qy        841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGTTCTTCAAG       900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3977 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGTTCTTCAAG     4036

Qy        901 ACCCTGCGGGCCGAGCAGACGACCCAGGAGGTGAAGAACTGGATGACGGACACCCTGCTG       960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4037 ACCCTGCGCCGCCGAGCAGACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG     4096

Qy        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCGCTTCTGGGCCCCGCAGC      1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4097 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCGCTTCTGGGCCCCGCAGC     4156

Qy       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCGGGCCACAAGGCCCGGGTG      1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4157 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCGGGCCACAAGGCCCGGGTG     4216

Qy       1081 CTGGCCGAGGCCATGAGCCAGGTGACCAACAGCGCCAACAACAACCAGATGGCAACTTCAAG      1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4217 CTGGCCGAGGCCATGAGCCAGGTGACCAACAGCGCCAACAACAACCAGATGGCAACTTCAAG     4276

Qy       1141 GGCCCCAAGCGGATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGGAAC      1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4277 GGCCCCCGGGCCATCGTCAAGTGCTTCAACTGCGGGCAAGGAGGGCCACATCGCCCGGCAAC     4336

Qy       1201 TGCGGCGGCCGCGCGCCGCAAGAAGGCTGCTGGAAGTGTGGCGCAGAAGCGCATCGAAG      1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4337 TGCGGCGGCCGCGCGCCGCAAGAAGGCTGCTGGAAGTGTGGCGCAGAAGCGCATCGAAG     4396

Qy       1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCGCAGAAGGGCCGC      1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4397 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCGCAGAAGGGCCGC     4456

Qy       1321 CCCGGCAACTTCCTGCAGAGCCGCCCGGAGCCCACCGCCCCGCCCGGAGAGCTTCCGC      1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4457 CCCGGCAACTTCCTGCAGAGCCGCCCGGAGCCCACCGCCCCGCCCGGAGAGCTTCCGC     4516

Qy       1381 TTCGAGGAGGACCACCCCCGGCCGCCCGGGCGAGGCCGGGCCGAGACCCTGACCAGC      1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4517 TTCGAGGAGGACCACCCCCGGCCGCCCGGGCGAGGAGGACCCGGAGACCCTGACCAGC     4576

Qy       1441 CTGAAGAGCCTGTTCGGCGAACGACCCCCTGAGCCCAGTAA      1479
              ||||||||||||||||||||||||||||||||||||| ||||
Db       4577 CTGAAGAGCCTGTTCGGCGAACGACCCCCTGAGCCCAATAA      4615
```

RESULT 10

```
ACA03522
ID   ACA03522 standard; DNA; 4716 BP.
XX
AC   ACA03522;
XX
DT   22-MAY-2003  (first entry)
XX
DE   Synthetic DNA encoding immunogenic HIV peptide #5.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
PN   WO2003004657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Pig 10; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation of, or for generating
CC   an immune response (e.g. a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
CC   human immunodeficiency virus immunogenic peptide
XX
SQ   Sequence 4716 BP; 1041 A; 1684 C; 1463 G; 528 T; 0 U; 0 Other;
```

Query Match      99.9%;   Score 1477.4;   DB 7;   Length 4716;
Best Local Similarity   99.9%;
Matches  1478;   Conservative    0;   Mismatches   1;   Indels    0;   Gaps    0;
Pred. No. 3.1e-185;

(DNA sequence alignment between Query (Qy) and Database (Db), positions 1–1479 vs 7–4702, aligned in consecutive 60-base rows with matching bars.)

```
Db   187  ATCCGCGCAGCTTCCACCCCTGCAGACCCGGCGGCGAAGGGCAGGGAGCTGAAGAGCCTGTTCAAC  246
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   241  ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTGCGCGACACCAAGGAGGCC         300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247  ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTGCGCGACACCAAGGAGGCC         306
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   301  CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGACGCAGGCC                 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307  CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGACGCAGGCC                 366
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   361  GCCGACAAGGGCAAGGTGAGCGAGCAGAACAAGAGCAAGCAGAAGAGCCAGATG             420
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367  GCCGACAAGGGCAAGGTGAGCGAGCAGAACAAGAGCAAGCAGAAGAGCCAGATG             426
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421  GTGCCACCAGGCCATCAGCCCCTGAAGCCTGGGTGAAGGTGATCGAGGAG                  480
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   427  GTGCCACCAGGCCATCAGCCCCTGAAGCCTGGGTGAAGGTGATCGAGGAG                  486
          |||||||||||||||||||||||||||||||||||||||||||||||||
Qy   481  GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGC                    540
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   487  GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGC                    546
          ||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541  GACCTGAACCACGATGTTGAACACCGTGGGCGGCCACCAGGCTGAAG                     600
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   547  GACCTGAACCACGATGTTGAACACCGTGGGCGGCCACCAGGCTGAAG                     606
          |||||||||||||||||||||||||||||||||||||||||||||||
Qy   601  GACACCATCAACCAGGAGGAGGCCGAGTGGGACCGGGTGCACCCGCCC                    660
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   607  GACACCATCAACCAGGAGGAGGCCGAGTGGGACCGGGTGCACCCGCCC                    666
          ||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661  ATCGCCCCCGGCCAGATGCGCGAGCCCCGGGGCAGCGACATCGCCGGCACCACCAGCACC        720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   667  ATCGCCCCCGGCCAGATGCGCGAGCCCCGGGGCAGCGACATCGCCGGCACCACCAGCACC        726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721  CTGCAGGAGCAGATCGGCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC        780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   727  CTGCAGGAGCAGATCGGCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC        786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781  AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC        840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   787  AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC        846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   841  CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGGTTCTTCAAG        900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   847  CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGGTTCTTCAAG        906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901  ACCCTGCGCGCGGAGCAGGCGACCCAGGAGGTGAAGAACTGGATGACCGAGACCCTGCTG        960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   907  ACCCTGCGCGCGGAGCAGGCGACCCAGGAGGTGAAGAACTGGATGACCGAGACCCTGCTG        966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   961  GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTCGGCCCTGGCGCCAGC       1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   967  GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTCGGCCCTGGCGCCAGC       1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1021  CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG       1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1027  CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG       1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1081  CTGGCCCGAGGCCGATGACCGCCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAGC     1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1087  CTGGCCCGAGGCCGATGACCGCCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAGC     1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1141  GGCCCCCCGCCGGCCATCGTCAAGCTTCAACTCGGGCAAGGAGGGGCCACATCGGCCGCAAC     1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1147  GGCCCCCCGCCGGCCATCGTCAAGCTTCAACTCGGGCAAGGAGGGGCCACATCGGCCGCAAC     1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1201  TGCCGCGCCCGCCAAGAGGCTGCGTGGAGTGGCGGCCAAGAAGGAGAGGGCCAGGATGAAG     1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1207  TGCCGCGCCCGCCAAGAGGCTGCGTGGAGTGGCGGCCAAGAAGGAGAGGGCCAGGATGAAG     1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1261  GACTGCACCGAGCGGCCAGGGCAACTTCCTGGGCCAAGGATCTGGGCCCAGCCCACAAGGGGCCGC   1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1267  GACTGCACCGAGCGGCCAGGGCAACTTCCTGGGCCAAGGATCTGGGCCCAGCCCACAAGGGGCCGC   1326
```

```
Qy  1321  CCCGGCCAACTTCCTGCAGAGAGCCGCCCCCAGGACCGCCCACCCGCCCCCCGCCGCGAGAGCTTCCGC  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1327  CCCGGCCAACTTCCTGCAGAGAGCCGCCCCCAGGACCGCCCACCCGCCCCCCGCCGCGAGAGCTTCCGC  1386
Qy  1381  TTCGGAGGAGGACCACCCCGGCCCCAGCAGCCCCCAGAAGCAGGAGGCAGGCGAGGACCCGACCAGC    1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1387  TTCGGAGGAGGACCACCCCGGCCCCAGCAGCCCCCAGAAGCAGGAGGCAGGCGAGGACCCGACCAGC    1446
Qy  1441  CTGAAGGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA                              1479
          ||||||||||||||||||||||||||||||||||
Db  1447  CTGAAGGAGCCTGTTCGGCAACGACCCCCTGAGCCAGAA                               1485
```

RESULT 11
ADC13238
ID   ADC13238 standard; DNA; 4716 BP.
XX
AC   ADC13238;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagProrinaRtumtTatRevNef_C SEQ ID NO 17.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002NO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-03498710P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 14; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intravenously, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 4716 BP; 1041 A; 1684 C; 1463 G; 528 T; 0 U; 0 Other;

Query Match      99.9%;  Score 1477.4;  DB 9;  Length 4716;
Best Local Similarity   99.9%;  Pred. No. 3.1e-185;
Matches 1478;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

```
Qy     1  ATGGGCGCCCGCAGCATCCTCCGCGGCAAGCTGACGGCCTGGGAGCGCATCCGC   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy      7 ATGGGCGCGCCCGGCAGCATCCTGCGCGGCGCCAAGCTGCTGGACGGGCATCCGC 66
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7 CTGCGCCCCCGGCGGCAAGAAGTGCTACTACATGATGAAGCACCTGGTGTGGGCCAGCAGCCGGCCGAG 66

Qy     61 CTGCGGAAGAAGTTCGCCCTGACACCCGGCCTGCTTGGAGACACCAGCAAGCAGATC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGCGGAAGAAGTTCGCCCTGACACCCGGCCTGCTTGGAGACACCAGCAAGCAGATC 120

Qy     67 CTGGCAGAAGTTCGCCCTGACACCCGGCCTGCTTGGAGACACCAGCAAGCAGATC 126
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 CTGGCAGAAGTTCGCCCTGACACCCGGCCTGCTTGGAGACACCAGCAAGCAGATC 126

Qy    121 CTGGGAGAAGTTCGCCCTGACACCCGGCCTGCTTGGAGACACCAGCAAGCAGATC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGGGAGAAGTTCGCCCTGACACCCGGCCTGCTTGGAGACACCAGCAAGCAGATC 180

Qy    127 CTGGAGAAGTTCGCCCTGACACCCGGCCTGCTTGGAGACACCAGCAAGCAGATC 186
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 186

Qy    181 ATCCCGCCAGCTGCACCCCGCCCTGCAGAACCCGGCGCCAGCAGGCAGCTTCAAC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATCCCGCCAGCTGCACCCCGCCCTGCAGAACCCGGCGCCAGCAGGCAGCTTCAAC 240

Qy    187 ATCCCGCCAGCTGCACCCCGCCCTGCAGAACCCGGCGCCAGCAGGCAGCTTCAAC 246
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187 ATCCCGCCAGCTGCACCCCGCCCTGCAGAACCCGGCGCCAGCAGGCAGCTTCAAC 246

Qy    241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGGGACACACAGGAGGCC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGGGACACACAGGAGGCC 300

Qy    247 ACCGTGGCCACCCTGTTACTGCGTGCACGAGAGAAGATCGAGGTCCGGGACACCAAGGAGGCC 306
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    247 ACCGTGGCCACCCTGTTACTGCGTGCACGAGAGAAGATCGAGGTCCGGGACACCAAGGAGGCC 306

Qy    301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAAGCAGAAGATCCAGCAGCAGGCC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAAGCAGAAGATCCAGCAGCAGGCC 360

Qy    307 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAAGCAGAAGATCCAGCAGCAGGCC 366
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    307 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAAGCAGAAGATCCAGCAGCAGGCC 366

Qy    361 GCCGACAGAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCCAGATG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCGACAGAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCCAGATG 420

Qy    367 GCCCGACAAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCCAGATG 426
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    367 GCCCGACAAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCCAGATG 426

Qy    421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480

Qy    427 GTGCACCAGGCCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 486
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    427 GTGCACCAGGCCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 486

Qy    481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540

Qy    487 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 546
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    487 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 546

Qy    541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600

Qy    547 GACCTGAACACGATGTTTGAACACCGTGGGCGGCCCACCAGGCCGCCCACCAGATGCTGAAG 606
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    547 GACCTGAACACGATGTTTGAACACCGTGGGCGGCCCACCAGGCCGCCCACCAGATGCTGAAG 606

Qy    601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGCTGCACCCCGTGCACGCCGGGCCCC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGCTGCACCCCGTGCACGCCGGGCCCC 660

Qy    607 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGCTGCACCCCGTGCACGCCCGGGCCCC 666
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    607 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGCTGCACCCCGTGCACGCCCGGGCCCC 666

Qy    661 ATCGCGCCCGGCCAGATGCGCGAGCCCCGGGGCAGCGACATCGCCGGCACCACCAGCACC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ATCGCGCCCGGCCAGATGCGCGAGCCCCGGGGCAGCGACATCGCCGGCACCACCAGCACC 720

Qy    667 ATCGCGCCCGGCCAGATGCGCGAGCCCCGGGGCAGCGACATCGCCGGCACCACCAGCACC 726
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    667 ATCGCGCCCGGCCAGATGCGCGAGCCCCGGGGCAGCGACATCGCCGGCACCACCAGCACC 726

Qy    721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780

Qy    727 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 786
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    727 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 786

Qy    781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840

Qy    787 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 846
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    787 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 846

Qy    841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 900

Qy    847 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 906
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    847 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 906

Qy    901 ACCCTGCGCGCCGAGCAGGCGGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ACCCTGCGCGCCGAGCAGGCGGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGC 960

Qy    907 ACCCTGCGCGCCGAGCAGGCGGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGC 966
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    907 ACCCTGCGCGCCGAGCAGGCGGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGC 966

Qy    961 GTGCAGAACGCCAACCCGACTGCAAGACCATCCTGGCGGCTCTCGGCCCGGCCAGC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GTGCAGAACGCCAACCCGACTGCAAGACCATCCTGGCGGCTCTCGGCCCGGCCAGC 1020

Qy    967 GTGCAGAACGCCAACCCGACTGCAAGACCATCCTGGCGGCTCTCGGCCCGGCCAGC 1026
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    967 GTGCAGAACGCCAACCCGACTGCAAGACCATCCTGGCGGCTCTCGGCCCGGCCAGC 1026

Qy   1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1080

Qy   1027 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1086
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1027 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1086

Qy   1081 CTGGCCGAGGCCGATGAGCCAGGCCAGCGTCAGCGTGATGAGCAGGCAGAGAGACAACAC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CTGGCCGAGGCCGATGAGCCAGGCCAGCGTCAGCGTGATGAGCAGGCAGAGAGACAACAC 1140

Qy   1087 CTGGCCGAGGCCGATGAGCCAGGCCAGCGTCAGCGTGATGAGCAGGCAGAGAGACAACAC 1146
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1087 CTGGCCGAGGCCGATGAGCCAGGCCAGCGTCAGCGTGATGAGCAGGCAGAGAGACAACTTCAAG 1146

Qy   1141 GGCCCCCGGCGCATCGTTCAAGTCGTTCAACTCGCGGCGCAGGAGGGCCACGCCCGCCAAC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GGCCCCCGGCGCATCGTTCAAGTCGTTCAACTCGCGGCGCAGGAGGGCCACGCCCGCCAAC 1200

Qy   1147 GGCCCCCGGCGCATCGTTCAAGTGCTTCAACTCGCGGCGCAGGAGGGCCACATCGGCCCGCCAAC 1206
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1147 GGCCCCCGGCGCATCGTTCAAGTGCTTCAACTCGCGGCGCAGGAGGGCCACATCGGCCCGCCAAC 1206

Qy   1201 TGCCGGCCCCCGCCCCCGCCCAAGGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TGCCGGCCCCCGCCCCCGCCCAAGGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260

Qy   1207 GACTGCACCGAGCGGCCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1266
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1207 GACTGCACCGAGCGGCCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1266

Qy   1261 GACTGCACCGAGCGGCCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GACTGCACCGAGCGGCCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320

Qy   1321 CCCGGCAACTTCCTGCAGAGCCGGCCCGAGCCCACCGCCCCACCCGAAGAGAGCTTCCGC 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CCCGGCAACTTCCTGCAGAGCCGGCCCGAGCCCACCGCCCCACCCGAAGAGAGCTTCCGC 1380

Qy   1327 CCCGGCAACTTCCTGCAGAGCCGGCCCGAGCCCACCGCCCCACCCGAAGAGAGCTTCCGC 1386
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1327 CCCGGCAACTTCCTGCAGAGCCGGCCCGAGCCCACCGCCCCACCCGAAGAGAGCTTCCGC 1386

Qy   1381 TTCGGAGGAGACCACCCCGCCCCCGCCCCGGAGCAGCAGGACGAGAGCAAGGACCCTGACCAGC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TTCGGAGGAGACCACCCCGCCCCCGCCCCGGAGCAGCAGGACGAGAGCAAGGACCCTGACCAGC 1440

Qy   1387 TTCGGAGGAGACCACCCCGCCCCCGCCCCGGAGCAGCAGGACGAGAGCAAGGACCCTGACCAGC 1446
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1387 TTCGGAGGAGACCACCCCGCCCCCGCCCCGGAGCAGCAGGACGAGAGCAAGGACCCTGACCAGC 1446

Qy   1441 CTGAAGAGCCTGTTCGGCAACGACACCCCCTGAGCCCAGTAA 1479
            ||||||||||||||||||||||||||||||||||||| ||
Db   1447 CTGAAGAGCCTGTTCGGCAACGACACCCCCTGAGCCCAGAAA 1485
```

RESULT 12
ACA03524

ID   ACA03524 standard; DNA; 2742 BP.
XX
AC   ACA03524;
XX
DT   22-MAY-2003  (first entry)
XX
DE   Synthetic DNA encoding immunogenic HIV peptide #7.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
XX
FN   WO200300465?-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic HIV polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Fig 12; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptide, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded protein are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular

CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation of, or for generating
CC   an immune response (e.g. a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
CC   human immunodeficiency virus immunogenic peptide
XX
SQ   Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 U; 0 Other;

Query Match             99.8%;  Score 1475.8;  DB 7;  Length 2742;
Best Local Similarity   99.8%;  Pred. No. 5.4e-185;
Matches 1477;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy          1 ATGGGCGCCCCCAGGGCCCCGCGGCGCGGCCCAAGCTGGGGGCGGCGGCCGATCCGGCGCCGCCATCCGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 ATGGGCGCCCCCAGGGCCCCGCGGCGCGGCCCAAGCTGGGGGCGGCGGCCGATCCGGCGCCGCCATCCGC 66

Qy         61 CTGGCGCCCCCGGGCGGCGGCGAAGAAGTGCTACATGTATGAAGCACCTGGTGCCAGCGCGCGCCAGCGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CTGGCGCCCCCGGGCGGCGGCGAAGAAGTGCTACATGTATGAAGCACCTGGTGCCAGCGCGCGCCAGCGAG 126

Qy        121 CTGGAGAAGTTCGCCCTTGGGCCGGGCTTGCTTGCTGTGGAGAACCCGGGCTTGCAAGCAGATC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CTGGAGAAGTTCGCCCTTGGGCCGGGCTTGCTTGCTGTGGAGAACCCGGGCTTGCAAGCAGATC 186

Qy        181 ATCCGGCCAGCTGCACCCCTTCCACCCCGGCCCTGCTGTCGAGACCCGGGCTGCAGAGCCTGGTTGTTCAAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ATCCGGCCAGCTGCACCCCTTCCACCCCGGCCCTGCTGTCGAGACCCGGGCTGCAGAGCCTGGTTGTTCAAC 246

Qy        241 ACCGTGGCGCCACCCTGTACTCGCGTCGCACGAGAAGATCGAGGTTCCGGACACCAAGAAGCAGGAGGCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ACCGTGGCGCCACCCTGTACTCGCGTCGCACGAGAAGATCGAGGTTCCGGACACCAAGAAGCAGGAGGCC 306

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGAGTGCCAGCAGAAGATCCACAGGCCCGAGGCC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGACAAGATCGAGGAGGAGCAGAACAAGAGTGCCAGCAGAAGATCCACAGGCCCGAGGCC 366

Qy        361 GCCGACAAGGGGCAAGGTGAGCCAGGAACTACCCCATCGTTGCACAGAACCTTGCAGCAGCGATG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GCCGACAAGGGGCAAGGTGAGCCAGGAACTACCCCATCGTTGCACAGAACCTTGCAGCAGCGATG 426

Qy        421 GTGCACAGGCCATCAGCCCCGCGACCCTGAAGCCCTGGGTGAAGGTGATCGAGGAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GTGCACAGGCCATCAGCCCCGCGACCCTGAAGCCCTGGGTGAAGGTGATCGAGGAAG 486

Qy        481 GCCTTCAGCCCCGGAGGTGATCCCCCATGTTCACCCGCCTGAGCGAGGGGCCACCCCCAG 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GCCTTCAGCCCCGGAGGTGATCCCCCATGTTCACCCGCCTGAGCGAGGGGCCACCCCCAG 546

Qy        541 GACCTGAACACGGATGTTGAACACCGTGGGCGGCCCCCCATCCAGATGCTGAAG 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACCTGAACACGGATGTTGAACACCGTGGGCGGCCCCCCATCCAGATGCTGAAG 606

Qy        601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGGTGCACCCGGTGCACGCCGGCCC 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGGTGCACCCGGTGCACGCCGGCCC 666

Qy        661 ATCGCCCCCGGCCGCAGATCGCGCACATCCCGGCGCACCCACCACAGCACC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 ATCGCCCCCGGCCGCAGATCGCGCACATCCCGGCGCACCCACCACAGCACC 726

Qy        721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCACCATCCGGTGGGCCGACATCTAC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCACCATCCGGTGGGCCGACATCTAC 786

Qy        781 AAGCGGTGTGATCATCCTGGGCGTGGGATCTGACAAGATCGGGAGTGTACAGCGCCCCGTGAGCGATC 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 AAGCGGTGGATCATCCTGGGCGTGGGATCTGACAAGATCGGGAGTGTACAGCGCCCCGTGAGCGATC 846

Qy        841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGGCTTCTTCAAG 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGGCTTCTTCAAG 906

Qy        901 ACCCTTGCGCGCGCCCGAGCAGAGCACCCCAGCGAGGTGGAGGTGAAGAACTGGATGACCGAGACACCCTGGCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 ACCCTTGCGCGCGCCCGAGCAGAGCACCCCAGCGAGGTGGAGGTGAAGAACTGGATGACCGAGACACCCTGGCTG 966

Qy        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTTCGCGCCCTGGGCCCGCCCCAGGCCAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTTCGCGCCCTGGGCCCGCCCCAGGCCAGC 1026

Qy       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1086

Qy       1081 CTGGCCGAGGCGATGAGCCAAGCCAACCACACCAGCGTGATGGTGCAGAAGAGCAACTTCAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 CTGGCCGAGGCGATGAGCCAAGCCAACCACACCAGCGTGATGGTGCAGAAGAGCAACTTCAAG 1146

Qy       1141 GGCCCCCGCCGCACATCGTTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACAATGCCGCCAAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 GGCCCCCGCCGCACATCGTTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACAATGCCGCCAAC 1206

Qy       1201 TGCCGCGCCCCCGCCGCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 TGCCGCGCCCCCGCCGCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1266

Qy       1261 GACTGCACCGAGCGCGCCAGCGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 GACTGCACCGAGCGCGCCAGCGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1326

Qy       1321 CCCGGCAACTTCCTGCAGAGCCGCCCGGAGCCCACCGCCCCGCCGGAGCTTCCGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 CCCGGCAACTTCCTGCAGAGCCGCCCGGAGCCCACCGCCCCGCCGGAGCTTCCGC 1386

Qy       1381 TTCGGAGGAGGACCACCCCCCGCCCGGCCAGGAGCCGGAGGCCAAGGACCCTGACCCAGC 1440
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1387 TTCGGAGGAGGACCACCCCCCGCCCGGCCAGGAGCCGGAGGCCAAGGACCCTGACCCAGC 1446

Qy       1441 CTGAAGAGCCTGTTCGGCAACGAACGACCCCCTGAGCCAGTAA 1479
              |||||||||||||||||||||||||||||||||||| ||
Db       1447 CTGAAGAGCCTGTTCGGCAACGAACGACCCCCTGAGCCAAGRA 1485

RESULT 13
ADC13241

ID   ADC13241   standard; DNA; 2742 BP.
XX
AC   ADC13241;
XX
DT   18-DEC-2003   (first entry)
XX
DE   DNA of HIV construct GagTatRevNef_C SEQ ID NO 20.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; de.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg BJ;
XX
DR   WPI; 2003-221593/21.

PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.

XX
PS   Disclosure; Fig 17; 301pp; English.
XX

CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 U; 0 Other;

Query Match   99.8%;   Score 1475.8;   DB 9;   Length 2742;
Best Local Similarity   99.8%;   Pred. No. 5.4e-185;
Matches 1477; Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

```
Qy        1 ATGGGCCCCGCCCAGCCATCCTCGCGCGCGGGGCGCGGGGCGCCATCCGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7 ATGGGGCCGGCCCAGCCATCCTGCGCGCGCAGCGTGGACGCCCATCCGC 66

Qy       61 CTGCGCCCCGGCGGCGCAAGAAGTCCTACATGATGATGAACGCCAGCCGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       67 CTGCGCCCCGGCGGCCAAGAAGTCCTACATGATGAACGCCAGCCCGGAG 126

Qy      121 CTGGAGAAGTTCGCCCTGTGAACCCCCGGCCGCTGCTTGGGAGACCAGCAAG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      127 CTGGAGAAGTTCGCCCTGCTTGGGAGACCACCTGGTTGTGGCCAGCGCGAG 186

Qy      181 ATCCGCCAGCTTCGCACCCGATCACCCGCCCCTGGTCTGTGAAGCAC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      187 ATCCGCCAGCTTCGCACCCCGCCCCTGGTCTGTGGGCCAAGGCAGATC 246

Qy      241 ACCGTGCGCCACCCTGTACTGCGTGCCACCGAGAGATGAGGTCCGGCAGCC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      247 ACCGTGGCCACCCGTACTGGGTGCCACCGAGAGATCGAGGTCCGGCACGC 306

Qy      301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGCAGCCGAGCC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      307 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGCAGCCGAGCC 366

Qy      361 GCCCGACAAGGGCAAGGTGAGCGAGGCCAGAACTACCCCATCGTGCACAGAAG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      367 GCCCGACAAGGGCAAGGTGAGCCCAGAACTACCCCATCGTGCAGAACTTGCAGAGCCCAGATG 426

Qy      421 GTGCACCAGGGCATCAGCCCCGACACCCGAGCGCTGGGTGAAGGTGATCGAGGAGAAG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      427 GTGCACCAGGGCCATCAGCCCCGACACCCGAACGGCCTGGGGTGAAGGTGATCGAGGAGATC 486

Qy      481 GCCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGGGACCCAG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      487 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGGGCCCC 546

Qy      541 GACTGAACGATGTTGAACACCGTGGGCGGCCATGCAGGATGTCTGAAG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      547 GACTGAACACGATGTTGAACACCGTGGGCGGCCATGCAGATGCCTGAAG 606

Qy      601 GACACCATCAACGAGGAGGCCGAGGTGGACAGGTTCACGCTGCACCCGGCC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      607 GACACCATCAACGAGGAGGCCGAGGTGGACAGGTTCACGCTGCACCCGGCC 666

Qy      661 ATCGGCCCCCGGCCAGATGCGGCGCAGCACATCGCCCGGCCACCAGCACC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      667 ATCGGCCCCCGGCCAGATGCGGCAGCACATCGCCGGGCCACCAGCACC 726

Qy      721 CTGGAGGAGAGAGAGATGCCCTTGCCAGGCGTGGACGGCCAGCAGCAACCCCCCATCCCCGTGGGGGGACATCTAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      727 CTGCAGGAGAGAGAGATGCCCTGATGACCACCGCAACCCCCCATCCCCGTGGGGGGACATCTAC 786

Qy      781 AAGCGGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCGGTGACCATC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      787 AAGCGGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCGGTGACCATC 846

Qy      841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      847 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 906

Qy      901 ACCCTGCGCGCCGAGCAGGCCACCCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      907 ACCCTGCGCGCCGAGCAGGCCACCCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 966

Qy      961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCTCTGGGCCCCGCTG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGGCTCTGGGCCCCGGCCAGC 1026

Qy     1021 CTGGAGGAGGATGAGGATGACCCGCCTGCCAGGGGCGGTGGCCGCCAGGCCCAGCGTG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1027 CTGGAGGAGGATGAGGATGACCCGGCCGGGCCTGCCCAGGGGCGTGGCCAGCGCCCAGGCCGTG 1086

Qy     1081 CTGGCCCGAGCGGATGAGCCGCCAACAACACACGCTGGCACAGAGAGCACAAGCCAACTTCAAG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1087 CTGGCCGGAGCGGATGAGCCGCCAACAACACACGCTGGGCCCCCGCAAGAGACACCAGGCAACTTCAAG 1146

Qy     1141 GGCCCCCGCCCGGCCCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATTGCCCGCAAC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1147 GGCCCCCGGGCCCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATTGCCCGCAAC 1206

Qy     1201 TGCCCGCCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCAGATGAAG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1207 TGCCCGCCCCCGCCCAAGGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGAAG 1266

Qy     1261 GACTGCACCGAGCGCCCAGCCCACTTCCTGGGCAAGATCTGGCCCCAGCCCAAGGGCGC 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1267 GACTGCACCGAGCGCCCAGCCCACTTCCTGGGCAAGATCTGGCCCAGCCCAAGGGGCGC 1326

Qy     1321 CCCGGCAACTTCCTGCAGAGCAGGCCCGAGCCCACCGCCCCCCCGAGAGCTTCCGC 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1327 CCCGGCAACTTCCTGCAGAGCAGGCCCGAGCCCACCGCCCCCCCCGAGAGCTTCCGC 1386

Qy     1381 TTCGAGGAGACCACCCCGCCCCCCAGCAGCAGGACCCTGACCAGC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1387 TTCGAGGAGACCACCCCGGCCCCCCAGCAGCAGGACCCTGACCAGC 1446

Qy     1441 CTGAAGAGCCTGTTCGGCAACGACCCCCCTGAGCCCAGTAA 1479
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1447 CTGAAGAGCCTGTTCGGCAACGACCCCCCTGAGCCCAAGAA 1485
```

RESULT 14
ADC13230
ID   ADC13230 standard; DNA; 3930 BP.
XX
AC   ADC13230;
XX
DT   18-DEC-2003   (first entry)
XX
DE   DNA of HIV construct GagComp1polmut_C SEQ ID NO 9.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO200300462O-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.

```
XX
PR   05-JUL-2001; 200lUS-0303192P.
PR   31-AUG-2001; 200lUS-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg BJ;
XX
DR   WPI; 2003-221593/21.
XX
PS   Disclosure; Fig 6; 301pp; English.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intranasally,
CC   subcutaneously, intradermally, transdermally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3930 BP; 890 A; 1365 C; 1214 G; 461 T; 0 U; 0 Other;

Query Match             99.8%;  Score 1475.8;  DB 9;  Length 3930;
Best Local Similarity   99.9%;
Matches 1477;  Conservative   0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 ATGGGGCCGCCGCCGCCAGCCATCTCTGCGCGCGGCGCAAAGCTGGACGCGCTGGGAGCGG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7 ATGGGGCCGCCGCCGCCAGCCATCTCTGCGCGCGGCGCAAAGCTGGACGCGCTGGGAGCGG 66

Qy       61 CTGCGCCCCCGGCGGCGCAAGAGTGCTCTACATCGTGTGGGCCCAGCGCCGCGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       67 CTGCGCCCCCGGCGGCGCAAGAGTGCTCTACATCGTGTGGGCCCAGCGCCGCGAG 126

Qy      121 CTTGGAGAAGTTCGCCCCTGAACCCCGGCGTCCTGCTGCTGGAGGGGTCGCAAGCAGATC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      127 CTTGGAGAAGTTCGCCCCTGAACCCCGGCGTCCTGCTGCTGGAGGGGTCGCAAGCAGATC 186

Qy      181 ATCCGCCAGCTGCACCCCCCGCCCTCGAGACCGGCGCAGGGAGGAGCCTGTTCAAC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      187 ATCCGCCAGCTGCACCCCCCGCCCTCGAGACCGGCGCAGGGAGGAGCCTGTTCAAC 246

Qy      241 ACCGTGGGCCACCCCTGTACTGCGGTGCACGAGAAGATCGGAGGTCCCGGACACCAAGAGGC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      247 ACCGTGGGCCACCCCTGTACTGCGGTGCACGAGAAGATCGGAGGTCCCGGACACCAAGAGGC 306

Qy      301 CTTGGACAAGATCGAGGAGGAGGCAGAGAACAAGCGTCCCGACAATCCAGCAGTGGCC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      307 CTTGGACAAGATCGAGGAGGAGGCAGAGAACAAGCGTCCCGACAATCCAGCAGTGGCC 366

Qy      361 GCCGACAAAGGGCAAAGGTCAGCAGAACTACCCCATCCCCTGGGTGAAGGTGAAGGAAG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      367 GCCGACAAAGGGCAAAGGTCAGCAGAACTACCCCATCCCCTGGGTGAAGGTGAAGGAAG 426

Qy      421 GTGCACCAGGCCATCAGCCCCGCACCCCGGGTGAACGCCTGGACATCGGCGGGAG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      427 GTGCACCAGGCCATCAGCCCCGCACCCCGGGTGAACGCCTGGACATCGGCGGGAG 486

Qy      481 GCCTTCAGCCCCGAGGTGATCCCAATGTTCACCGCCCTGAGCGAGGGCGGGGGCCACCCCCCCG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      487 GCCTTCAGCCCCGAGGTGATCCCAATGTTCACCGCCCTGAGCGAGGGCGGGGGCCACCCCCCCG 546

Qy      541 GACCTGAACACGGATGTTGAACACACCGTGGGCCGGCCGGCCATGCAGATGCTGAAG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      547 GACCTGAACACGGATGTTGAACACACCGTGGGCCGGCCGGCCATGCAGATGCTGAAG 606

Qy      601 GACACCATCAACGAGGAGGCGGCCGCCGAGTGGGACCGGGCCCGGTGCACGCCCGCCCCC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      607 GACACCATCAACGAGGAGGCGGCCGCCGAGTGGGACCGGGCCCGGTGCACGCCCGCCCCC 666

Qy      661 ATCGCCCCCGGCCAGATGCGCGAGCCGCGCGGCAGCGACATCGCCGGACCACCAGCAC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      667 ATCGCCCCCGGCCAGATGCGCGAGCCGCGCGGCAGCGACATCGCCGGACCACCAGCAC 726

Qy      721 CTGGACATCAAGGCAGGAGCCCAGATGCGGGAACAAGAACCCCCCATCCCGTGGGCGGGCATCTAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      727 CTGGACATCAAGGCAGGAGCCCAGATGCGGGAACAAGAACCCCCCATCCCGTGGGCGGGCATCTAC 786

Qy      781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      787 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 846

Qy      841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      847 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 906

Qy      901 ACCCTGCGCGGGCCCCGGAGCGAGGCCATCAAGGAGGTGAAGGAACTGGAATCACCACCCTGCTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      907 ACCCTGCGCGGGCCCCGGAGCGAGGCCATCAAGGAGGTGAAGGAACTGGAATCACCACCCTGCTG 966

Qy      961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCCTCGGGCCCGGGCCCAG 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCCTCGGGCCCGGGCCCAG 1026

Qy     1021 CTGGAGGAGGATGATGACCGCCGTCGCCAGGGGCGGCTGGGCGCCCAGCCCACAAGGCCCGGGTG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1027 CTGGAGGAGGATGATGACCGCCGTCGCCAGGGGCGGCTGGGCGCCCAGCCCACAAGGCCCGGGTG 1086

Qy     1081 CTGGCCGGAGCGGATGAGCCAGGCCAGGCCCAACAACCAGCGTGATGATGCAGAAGAGGCAACTTCAAG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1087 CTGGCCGGAGCGGATGAGCCAGGCCAGGCCCAACAACCAGCGTGATGATGCAGAAGAGGCAACTTCAAG 1146

Qy     1141 GGCCCCGGGGGCCCATCGTCAAGTGCTTCAACTGCGGCAAGGTGGGCCCAGGCCAAGGCCCCAAC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1147 GGCCCCGGGGGCCCATCGTCAAGTGCTTCAACTGCGGCAAGGTGGGCCCAGGCCAAGGCCCCAAC 1206

Qy     1201 TGCCGCGGCCCCCGCAAGAAGGGCTGCTGGAAGTGGGGCAAGAGGCCCACCGAGATGAAG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1207 TGCCGCGGCCCCCGCAAGAAGGGCTGCTGGAAGTGGGGCAAGAGGCCCACCGAGATGAAG 1266

Qy     1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCAACAGGGCCGC 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1267 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCAACAGGGCCGC 1326

Qy     1321 CCCGGGCCAACTTCCTGCAAGAGGCGGCCGCCACCGCCCGAGAGAGAGCTTCCGC 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1327 CCCGGGCCAACTTCCTGCAAGAGGCGGCCGCCACCGCCCGAGAGAGAGCTTCCGC 1386

Qy     1381 TTCGAGGAGACCACCCCCGCCCGGCCGAGAACGCAGGAAGGCAGACCCCTGACCAGC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1387 TTCGAGGAGACCACCCCCGCCCGGCCGAGAACGCAGGAAGGCAGACCCCTGACCAGC 1446

Qy     1441 CTGAAGAGCCTGTTCGGCAACGACGACCCCCTGAGCCAAGA 1485
            |||||||||||||||||||||||||||||||||||| ||
Db     1447 CTGAAGAGCCTGTTCGGCAACGACGACCCCCTGAGCCAAGA 1485

RESULT 15
ID   ADC13231;
AC   ADC13231;
XX
ID   ADC13231 standard; DNA; 3930 BP.
XX
DT   18-DEC-2003 (first entry)
XX
```

DE   DNA of HIV construct GagComplPolmutAtt_C SEQ ID NO 10.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO200300420-A2.
XX
PD   16-JAN-2003.
XX
PP   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J,  Barnett SW,  Lian Y,  Engelbrecht S,  Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunisation, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 7; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3930 BP; 889 A; 1365 C; 1214 G; 462 T; 0 U; 0 Other;

Query Match     99.8%;  Score 1475.8;  DB 9;  Length 3930;
Best Local Similarity   99.9%;  Pred. No. 5.1e-185;
Matches 1477;  Conservative    0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy          1 ATGGGCGCGCCCGCAGCCATCCTGCGCGGCGGCCAAGCTTGGACGCGCTGGGAGCGCATCCGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 ATGGGCGCGCCCGCCAAGCTTGCGCGGCGGCGGCCAAGCTTGGACGCGCATCCGC 66

Qy         61 CTGCGCCCCCGGCCCGGCGGCAAGAAGTGCTACATGGTGAAGGACGCCACCGAGGGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CTGCGCCCCCGGCCCGGCGGCAAGAAGTGCTACATGGTGAAGGACGCCACCGAGGGAG 126

Qy        121 CTGGAGAAGTTCGCCCTGAACCCCGGCCCTGTGGAGAACCGACGGGCTCAAAGCAGATC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CTGGAGAAGTTCGCCCTGAACCCCGGCCCTGTGGAGAACCGACGGGCTCAAAGCAGATC 186

Qy        181 ATCCGCCAGCTGCACCCCCAGCCCCGGCCGCAGGGAGCTGAAGAGACGCCTGTTCAAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ATCCGCCAGCTGCACCCCCAGCCCCGGCCGCAGGGAGCTGAAGAGACGCCTGTTCAAC 246

Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAGGATCGAGGTCGGACACACCAAGGAGGCC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ACCGTGGCCACCCTGTACTGCGTGCACGAGAGGATCGAGGTCGGACACACCAAGGAGGCC 306

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACGCAGGCCGAGGCC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACGCAGGCCGAGGCC 366

Qy        361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCATCGTCGTAGAACCTGCAGGGCCAGATG 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCATCGTCGTAGAACCTGCAGGGCCAGATG 426

Qy        421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 486

Qy        481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 546

Qy        541 GACCTGAACACGATGTTGAACACCGTGGGGGCCCCACCAGCCGCCGCCACCGCCATGCTGAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACCTGAACACGATGTTGAACACCGTGGGGGCCCCACCAGCCGCCGCCACCGCCATGCTGAAG 606

Qy        601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGGTGCACCCCGTGCACGCCGGGCCCC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGGTGCACCCCGTGCACGCCGGGCCCC 666

Qy        661 ATCGCCCCCGGGCCAGATGCGCGAGCCCCGGCCAGCGACGATGGGCCGGCCACCCAGCACC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 ATCGCCCCCGGGCCAGATGCGCGAGCCCCGGCCAGCGACGATGGGCCGGCCACCCAGCACC 726

Qy        721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCACCCATCCCCGTGGGCGACATCTAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCACCCATCCCCGTGGGCGACATCTAC 786

Qy        781 AAGCGGGTGGGAATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 AAGCGGGTGGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 846

Qy        841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 CTGGACATCAAGCAGGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 906

Qy        901 ACCCTGCGCGCCGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 ACCCTGCGCGCCGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 966

Qy        961 GTCCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCCTCTGGCCCCGGCCAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 GTCCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCCTCTGGCCCCGGCCAGC 1026

Qy       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAAGGCCCGTG 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAAGGCCCGTG 1086

Qy       1081 CTGGCCCAGGGCCAGATGAGGGACCCCAACACCAGGGCGTGGGCGCCCGGCTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 CTGGCCCAGGGCCAGATGAGGGACCCCAACACCAGGGCGTGGGCGCCCGGCTG 1146

Qy       1141 GGCCCCGGCGCGCCCATCGGTCAAGAGTGCTTCAACTGCGGCAAGGAGGCCCACATCGCCCCGCAAC 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 GGCCCCGGCGCGCCCATCGGTCAAGAGTGCTTCAACTGCGGCAAGGAGGCCCACATCGCCCCGCAAC 1206

Qy       1201 TGCCGGCGCGCCCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 TGCCGGCGCGCCCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1266

Qy       1261 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCCCCAAGATCTGGCCCAGCAACAAGGGGCGC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 GACTGCACCGAGCGGCAGGCCAACTTCCTGGGCCCCAAGATCTGGCCCAGCAACAAGGGGCGC 1326

Qy       1321 CCCGGCAACTTCCTGCAGAGCCGCCCGGAGCCCACCGCCCCGCCCGAGGACTTCCGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 CCCGGCAACTTCCTGCAGAGCCGCCCGGAGCCCACCGCCCCGCCCGAGGACTTCCGC 1386

Qy       1381 TTCGAGGAGACCACCCCCGCCCCCAAGCAGGAGCAGGAAGACCCTGACCAGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1387 TTCGAGGAGACCACCCCCGCCCCCAAGCAGGAGCAGGAAGACCCTGACCAGC 1446

Qy       1441 CTGAAGAGCCTGTTCGGCAACGACACCCCCTGAGCCAGTAA 1479

Db       1447  CTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAAGAA  1485
              |||||||||||||||||||||||||||||||||||| ||

Search completed: May 28, 2004, 09:49:09
Job time : 689.39 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 28, 2004, 09:05:41 ; Search time 6077.87 Seconds
(without alignments)
10547.179 Million cell updates/sec

Title:          US-09-475-704A-3
Perfect score:  1479
Sequence:       1 atggggcccgcgccagcat..........acgacccctgagccagtaa 1479
Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched:       3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters:       6940544
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  GenRmbl1:*

```
1:  gb_ba:*
2:  gb_htg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_vr:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

---

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1479 | 100.0 | 1479 | 6 | AX455887 | AX455887 Sequence |
| 2 | 1479 | 100.0 | 1479 | 6 | AX468543 | AX468543 Sequence |
| 3 | 1473.4 | 99.6 | 1479 | 6 | AX468547 | AX468547 Sequence |
| 4 | 1463 | 98.9 | 1479 | 6 | AX455904 | AX455904 Sequence |
| 5 | 1328.2 | 89.8 | 1485 | 12 | AY181195 | AY181195 Synthetic |
| 6 | 1288.8 | 87.1 | 1491 | 6 | AX455983 | AX455983 Sequence |
| 7 | 1288 | 87.1 | 1494 | 6 | AX455935 | AX455935 Sequence |
| 8 | 1276.8 | 86.3 | 1509 | 6 | AX455888 | AX455888 Sequence |
| 9 | 1275.2 | 86.2 | 1509 | 6 | AX468544 | AX468544 Sequence |
| 10 | 1273.6 | 86.1 | 1509 | 6 | AX468548 | AX468548 Sequence |
| 11 | 1261.2 | 85.2 | 1509 | 6 | AX455905 | AX455905 Sequence |
| 12 | 1221.2 | 82.6 | 9166 | 6 | AX427930 | AX427930 Sequence |
| 13 | 1206 | 81.5 | 9788 | 6 | AX427936 | AX427936 Sequence |
| 14 | 1205.4 | 81.5 | 4288 | 12 | AX149648 | AX149648 Sequence |
| 15 | 1197.8 | 81.0 | 1509 | 12 | AP201927 | AP201927 Synthetic |
| 16 | 1197.8 | 81.0 | 1515 | 6 | BD263637 | BD263637 Improved |
| 17 | 1197.8 | 81.0 | 1515 | 6 | AR373320 | AR373320 Sequence |
| 18 | 1197.8 | 81.0 | 4472 | 6 | BD263697 | BD263697 Improved |
| 19 | 1197.8 | 81.0 | 4472 | 6 | AR373380 | AR373380 Sequence |
| 20 | 1197.8 | 81.0 | 4608 | 6 | BD263698 | BD263698 Improved |
| 21 | 1197.8 | 81.0 | 4608 | 6 | AR373381 | AR373381 Sequence |
| 22 | 1197.8 | 81.0 | 4689 | 6 | BD263696 | BD263696 Improved |
| 23 | 1197.8 | 81.0 | 4689 | 6 | AR373379 | AR373379 Sequence |
| 24 | 1197.8 | 81.0 | 4766 | 6 | BD263695 | BD263695 Improved |
| 25 | 1197.8 | 81.0 | 4766 | 6 | AR373378 | AR373378 Sequence |
| 26 | 1195.8 | 80.9 | 2031 | 6 | BD263640 | BD263640 Improved |
| 27 | 1195.8 | 80.9 | 2031 | 6 | AR373323 | AR373323 Sequence |
| 28 | 1159.6 | 78.4 | 1548 | 12 | AF287354 | AF287354 Synthetic |
| 29 | 1134.4 | 76.7 | 1539 | 6 | AX306431 | AX306431 Sequence |
| 30 | 1131.8 | 76.5 | 6438 | 6 | AX427919 | AX427919 Sequence |
| 31 | 1131.8 | 76.5 | 9194 | 6 | AX427925 | AX427925 Sequence |
| 32 | 1131.8 | 76.5 | 9194 | 6 | AX427926 | AX427926 Sequence |
| 33 | 1131.8 | 76.5 | 12411 | 6 | AX427927 | AX427927 Sequence |
| 34 | 1120.4 | 75.8 | 9169 | 6 | AX427931 | AX427931 Sequence |
| 35 | 1116.6 | 75.5 | 9170 | 6 | AX427928 | AX427928 Sequence |
| 36 | 1116.6 | 75.5 | 9189 | 6 | AX427921 | AX427921 Sequence |
| 37 | 1116.6 | 75.5 | 9792 | 6 | AX427932 | AX427932 Sequence |
| 38 | 1114.6 | 75.4 | 1847 | 12 | AF202464 | AF202464 Synthetic |
| 39 | 1114.6 | 75.4 | 1847 | 12 | AF202465 | AF202465 Synthetic |
| 40 | 1114.6 | 75.4 | 1853 | 6 | BD263638 | BD263638 Improved |
| 41 | 1114.6 | 75.4 | 1853 | 6 | AR373321 | AR373321 Sequence |
| 42 | 1114.6 | 75.4 | 1865 | 6 | BD263700 | BD263700 Improved |
| 43 | 1114.6 | 75.4 | 1865 | 6 | BD263701 | BD263701 Improved |
| 44 | 1114.6 | 75.4 | 1865 | 6 | AR373383 | AR373383 Sequence |
| 45 | 1114.6 | 75.4 | 1865 | 6 | AR373384 | AR373384 Sequence |

ALIGNMENTS

RESULT 1
AX455887

```
LOCUS       AX455887    1479 bp   DNA   linear   PAT 06-JUL-2002
DEFINITION  Sequence 3 from Patent WO0204493.
ACCESSION   AX455887
VERSION     AX455887.1  GI:21714880
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 3 17-JAN-2002;
```

CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

FEATURES    Location/Qualifiers
source    1..1479
    /organism="synthetic construct"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32630"
    /note="synthetic Gag of HIV strain AP110965"

ORIGIN

Query Match   100.0%;   Score 1479;   DB 6;   Length 1479;
Best Local Similarity   100.0%;   Pred. No. 4.1e-151;
Matches 1479;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

(DNA sequence alignment between Query (Qy) and Database (Db), positions 1–1479, with identity match lines throughout.)

RESULT 2
LOCUS    AX468543    1479 bp    DNA    linear    PAT 16-JUL-2002
DEFINITION   Sequence 63 from Patent WO0226209.
ACCESSION   AX468543
VERSION    AX468543.1    GI:21901373
KEYWORDS    .
SOURCE    Human immunodeficiency virus 1 (HIV-1)
ORGANISM   Human immunodeficiency virus 1
   Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
   lentivirus group.
REFERENCE   1
AUTHORS    O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
   Singh,M., Ulmer,J. and Dubensky,T.W.
TITLE    Microparticles for delivery of the heterologous nucleic acids
JOURNAL    Patent: WO 0226209-A 63 04-APR-2002;
   CHIRON CORPORATION (US)
FEATURES    Location/Qualifiers
source    1..1479
    /organism="Human immunodeficiency virus 1"
    /mol_type="unassigned DNA"
    /db_xref="taxon:11676"
ORIGIN

Query Match   100.0%;   Score 1479;   DB 6;   Length 1479;
Best Local Similarity   100.0%;   Pred. No. 4.1e-151;
Matches 1479;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

(DNA sequence alignment between Query (Qy) and Database (Db) continues.)

RESULT 3
LOCUS       AX468547              1479 bp    DNA        linear   PAT 16-JUL-2002
DEFINITION  Sequence 67 from Patent WO0226209.
ACCESSION   AX468547
VERSION     AX468547.1  GI:21901377
KEYWORDS    .
SOURCE      Human immunodeficiency virus 1 (HIV-1)
ORGANISM    Human immunodeficiency virus 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1
AUTHORS     O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
            Singh,M., Ulmer,J. and Dubensky,T.W.
TITLE       Microparticles for delivery of the heterologous nucleic acids
JOURNAL     Patent: WO 0226209-A 67 04-APR-2002;
            CHIRON CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1479
                     /organism="Human immunodeficiency virus 1"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11676"
ORIGIN

Query Match     99.6%;  Score 1473.4;  DB 6;  Length 1479;
Best Local Similarity   99.3%;  Pred. No. 1.6e-150;
Matches 1468;  Conservative 10;  Mismatches 1;  Indels 0;  Gaps 0;

Qy          1 ATGGGGCGCCCGGCCCAGCCATCCTCCGCGCGCGCGAAGCTGGAGCCGCTGGGAGCGATCCGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGGCGCCCGGCCCAGCCATCCTCCGCGCGCGCGAAGCTGGAGCCGCTGGGAGCGATC 60

Qy         61 CTGCGGCCCCGGCCCCGCGGCCCCGGCCCCGCAAGAGTGCTGGGCCCAGCCGCGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCGGCCCCGGCCCCGCGGCCCCGGCCCCGCAAGAGTGCTGGGCCCAGCCGCGAG 120

```
RESULT 4
AX455904
LOCUS       AX455904               1479 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 20 from Patent WO0204493.
ACCESSION   AX455904
VERSION     AX455904.1  GI:21714896
KEYWORDS
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 20 17-JAN-2002;
            CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..1479
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="synthetic Gag coding sequence of HIV strain
                     AF110965"
ORIGIN

Query Match     98.9%;  Score 1463;  DB 6;  Length 1479;
Best Local Similarity   99.3%;
Matches 1469;  Conservative   0;  Mismatches  10;  Indels   0;  Gaps   0;
```

The remainder of the page is a nucleotide sequence alignment between a Query (Qy) and Database (Db) sequence, printed in two columns of 60-base lines with a match bar ("||||") between each pair.

```
SOURCE      synthetic construct
KEYWORDS    .
VERSION     AY181195.1  GI:37413992
ACCESSION   AY181195
DEFINITION  Synthetic construct HIV-1-derived gag protein (gag) gene, complete
LOCUS       AY181195              1485 bp    DNA     linear   SYN 02-OCT-2003
            cds.
RESULT 5
```

```
ORGANISM    synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1485)
AUTHORS     Gao,F., Li,Y., Decker,J.M., Peyerl,F.W., Bibollet-Ruche,P.,
            Rodenburg,C.M., Chen,Y., Shaw,D.R., Allen,S., Musonda,R.,
            Shaw,G.M., Zajac,A.J., Letvin,N. and Hahn,B.H.
TITLE       Codon usage optimization of HIV type 1 subtype C gag, pol, env, and
            nef genes: in vitro expression and immune responses in
            DNA-vaccinated mice
JOURNAL     AIDS Res. Hum. Retroviruses 19 (9), 817-823 (2003)
REFERENCE   2  (bases 1 to 1485)
AUTHORS     Gao,F., Li,Y., Decker,J.M., Bibollet-Ruche,F., Rodenburg,C.M.,
            Chen,Y., Shaw,D.R., Allen,S., Musonda,R., Shaw,G.M., Zajac,A.J. and
            Hahn,B.H.
TITLE       Direct Submission
JOURNAL     Submitted (16-NOV-2002) Medicine, Duke University Medical Center,
            112 RPIII, Lasalle Street, Durham, NC 27710, USA
FEATURES             Location/Qualifiers
     source          1..1485
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="codon usage optimization of HIV-1 96ZM651"
     gene            1..1485
                     /gene="gag"
     CDS             1..1485
                     /gene="gag"
                     /codon_start=1
                     /transl_table=11
                     /product="gag protein"
                     /protein_id="AAO65393.1"
                     /db_xref="GI:37413993"
                     /translation="MGARASILRGGKLDKWEKIRLRPGGKKRYMIKHLVWASRELERF
                     ALNPGLLETSEGCKQIMKQLQPALQTGTEELRSLYNTVATIVYCVHSGVEVRDTKEALD
                     RIEEEQNKIQQKTQQAKAADKVSQNYPIVQNLQGQMVHQKLSPRTLNAWVKVIEE
                     KAFSPEVIPMFTALSEGATPQDLNTMLNTVGGHQAAMQLKDTINEEAAEWDRLHPVH
                     AGPIAPGQMREPRGSDIAGTTSTLQEQIAWMTSNPPIPVGDIYKRWIILGLNKIVRMY
                     SPVSILDIKQSFKEPFRDYVDRFFKTLRAEQATQEVKNWMTDTLLVQNANPDCKTILK
                     ALGPGATLEEMMTACQGVGGPSHKARVLAEAMSQTNSVNILMQRSNFKGNKRMVKCFN
                     CGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTEROANFLGKIWPSHKGRPGNPLQNR
                     PEPTAPPAESFRFEETTPAPKQESKDREALTSLKSLFGSDPLSQ"
ORIGIN
```

```
Query Match      89.8%;    Score 1328.2; DB 12;  Length 1485;
Best Local Similarity   94.3%;    Pred. No. 8e-135;
Matches 1401; Conservative    0; Mismatches  78;  Indels    6;  Gaps    2;
```

```
Qy        1 ATGGGCGCGCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACCGCCTGGACAAGTGGCGC 60
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||   |||||
Db        1 ATGGGCGCGCCGCGCCAGCGATCCTGCGCGGCGGCAAGCTGGACAAGTGGGAGAAGATCCGC 60

Qy       61 CTGCGCCCCGGCGGCAAGAAGCGCTACATGATCAAGCACCTGGTGTGGGCCAGCCGCGAG 120
            || |||||||||||||||||||||||||||||||| ||||||| ||| |||||||||||||
Db       61 CTGCGCCCCGGCGGCAAGAAGCGCTACATGATCAAGCACCTGGTGTGGGCCAGCCGCGAG 120

Qy      121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAAGGCTGCAAGCAGATC 180
            ||||||         |||||||||||||||||||||| ||||||||||||||||||||||
Db      121 CTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAAGGCTGCAAGCAGATC 180

Qy      181 ATCCGCCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTTCAAC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db      181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC 240

Qy      241 ACCGTGGCCACCATCGTGTACTGCGTGCACAGCGGCGTGGAGGTGCGCGACACCAAGGAGGCC 300
            |||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db      241 ACCGTGGCCACCCTGTACTGCGTGCACAGCGGCGTGGAGGTGCGCGACACCAAGGAGGCC 300

Qy      301 CTGGACAAGATCGAGGAGGAGCAGAACAAGATCCAGCAGAAAACCCAG 360
            ||||||                     |||     GCCAG
Db      301 CTGGACAGCGCCATCGAGGAGGAGCAGAACAAGATCCAGCAGAAAACCCAG 357

Qy      358 GCCGCGCGACAAGGGCCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAG 417
            |||||  |||||||||||
```

```
Qy      541 GACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGCTGAAGG 600
            |||||   |||
Db      541 GACACCAATCAACGAGAGCCCGAGTGGTCGGACCGGCTGCACCCGCCCACCAGG

Qy      601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGCTGCACCCCGTGCACGCCGGCCCC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGCTGCACCCCGTGCACGCCGGCCCC 660

Qy      661 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 720

Qy      721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780
            |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db      721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780

Qy      781 AAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACAGCCCCGTGAGCATC 840
            |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      781 AAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACAGCCCCGTGAGCATC 840

Qy      841 CTGGACATCAAGCAGAGCTTCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 CTGGACATCAAGCAGAGCTTCAAGGAGCCCTTCCGGGACTACGTGGACCGCTTCTTCAAG 900

Qy      901 ACCCTGCGCGCGGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      901 ACCCTGCGCGCGGAGCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960

Qy      961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGGGCCCCGGCGCCACCCTG 1020
            |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db      961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGGGCCCCGGCGCCACCCTG 1020

Qy     1021 GAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080

Qy     1081 CTGGCCGAGGCCATGAGCCAGACCAACAGCGTGATGATGATGCAGCGGAGCAACTTCAAG 1140
            ||||||||||||||||||||| |||||||||| |||||||||||||||||||||||||||
Db     1081 CTGGCCGAGGCCATGAGCCAGACCAACAGCGTGATGATGATGCAGCGGAGCAACTTCAAG 1140

Qy     1141 GGCCCCCGCCCGCCATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGGCCCCGAAC 1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 GGCAAGAACCGCATGGTGAAGTGCTTCAACTGCGGCAAGGAGGGGCACATCGCCCGGAAC 1200

Qy     1201 TGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 TGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260

Qy     1261 GACTGCACCGAGCGCGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGCCGC 1320
            ||||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||||
Db     1261 GACTGCACCGAGCGCGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGCCGC 1320

Qy     1321 CCCGGCAACTTCCTGCAGAACCGCCCCGAGCCCACCGCCCCGCCGAGAGCTTCCGG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 CCCGGCAACCCCCTGCAGAACCGCCCCGAGCCCACCGCCCCCCCCGCCGAGAGCTTCCGG 1380

Qy     1381 TTCGAGGAGACCACCCCCGCCCCCAAGCAGGAGAGCAAGGACCGCGAGGCCCTGACCAGC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 TTCGAGGAGACCACCCCCGCCCCCAAGCAGGAGAGCAAGGACCGCGAGGCCCTGACCAGC 1440

Qy     1441 CTGAAGAGCCTGTTCGGCAGCGACCCCCTGAGCCAGTAA 1479
            ||||||||||||||||||||||||||||||||||||||||
Db     1441 CTGAAGAGCCTGTTCGGCAGCGACCCCCTGAGCCAGTAA 1479
```

RESULT 6

AX455983

```
LOCUS       AX455983                1491 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 99 from Patent WO0204493.
ACCESSION   AX455983
VERSION     AX455983.1  GI:21714967
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 99 17-JAN-2002;
            CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..1491
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Gag_TV2_C_ZAopt"
ORIGIN

Query Match           87.1%;  Score 1288.8;  DB 6;  Length 1491;
Best Local Similarity 93.0%;  Pred. No. 1.5e-130;
Matches 1389;  Conservative 0;  Mismatches 87;  Indels 18;  Gaps 3;
```

[Sequence alignment region: paired Qy (query) and Db (database) nucleotide rows with match bars, spanning query positions 1–1485 and database positions 1–1491.]

RESULT 7

LOCUS       AX455935       1494 bp    DNA     linear     PAT 06-JUL-2002
DEFINITION   Sequence 51 from Patent WO0204493.
ACCESSION    AX455935
VERSION      AX455935.1    GI:21714919
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
          artificial sequences.
REFERENCE    1
  AUTHORS    zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE      Polynucleotides encoding antigenic hiv type c polypeptides,
          polypeptides and uses thereof
  JOURNAL    Patent: WO 0204493-A 51 17-JAN-2002;

FEATURES                 Location/Qualifiers
     source           1..1494
                   CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
                   /organism="synthetic construct"
                   /mol_type="unassigned DNA"
                   /db_xref="taxon:32630"
                   /note="HIV Type C Gag optimized"

ORIGIN

Query Match    Best Local Similarity   87.1%;    Score 1288;    DB 6;    Length 1494;
Matches 1393;   Conservative   92.9%;   Pred. No. 1.8e-130;   Indels 27;   Gaps 3;
Conservative   0;    Mismatches   80;

```
Qy    661 ATCGCCCCCCGGCCAGATGCCGGAGCCCCGGCGCAGGGACAATGGCGCCGGCGGCCACCAGCACC 720
              |||||||||||||||||||||||||||||||||||||||||||          |||||||
Db    658 GTGGCCCCCGGCCAGATGCCGGAGCCCCGGCGCAGGGACA            CCACCAGCATC 717

Qy    721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCGGTGGGGGGACATCTAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    718 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCGGTGGGGGGACATCTAC 777

Qy    781 AAGGCGGTGGGATCATCCTGGGCGTGCACCAGGCGGTGGCCGGCTGAGATCGATC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    778 AAGGCGCTGGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACAGCCCGATC 837

Qy    841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTCGACCGGTTCTTCAAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    838 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTCGACCGGTTCTTCAAG 897

Qy    901 ACCCTGCGCGCCGAGCAGGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 ACCTTGCGCGCCGAGCAGGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 957

Qy    961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGTG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGTG 1017

Qy   1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAAGGCCCGCGTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1018 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAAGGCCCGCGTG 1077

Qy   1081 CTGGCCGAGGCGGATGAGCCAGGC-----CAACACCAGCGGCTGATGATGCAGAAGAGCAACTTC 1137
              |||||||||||||||||||||     |||||||||||||||||||||||||||||
Db   1078 CTGGCAGAACGCCATGAGCGCCAACAACACCAGGGTGATGATCCAGAAGAGCAACTTC 1137

Qy   1138 AAGGGGCCCCCCCGGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGC 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1138 AAGGGGCCCCCCCGGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGC 1197

Qy   1198 AACTGCCGCGCCCCCCAAGAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1198 AACTGCCGCGCCCCCCAAGAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257

Qy   1258 AAGGACTGCACGGAGCGGCAGGCCAACTTCCTGGGCCAGAGCCGGCCAGCCAAGGGGC 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1258 AAGGACTGCACGGAGCGGCAGGCCAACTTCCTGGGCCAGAGCCGGCCAGCCAAGGGGC 1317

Qy   1318 CGCCCCGCCCAACTTCCTGCAGAGCCGCCCACCGCCCAGCCGGAGCGGGGGCCCACC 1377
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1318 CGCCCCGGCCAACTTCCTGCAGAGCCGCCCACCGCCCAGCCGGAGCGGGGGCCCACC 1377

Qy   1378 CGCTTCGGAGG---------AGACCACCCCGCCGAGCAGAAGCAGGAGCAGGAGGACCGC 1425
              |||||||||            |||||||||||||||||||||||||||||||||||||
Db   1378 GCCCCCCCGCCGAGAGCTTCAAGTTCGAAGAGACCACCCCGGCCCCAAGGCGGACCGC 1437

Qy   1426 GAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAGGGACCCCCTGAGCCAGTAA 1479
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1438 GAGCCCCTGACCAGCCTGAAGAGCCTGTTCGGCAGGGACCCCCTGAGCCAGTAA 1491
```

```
Qy     901 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
           ||||||||||||||||||  ||||||||| ||||||||||||||||||||||||||||||
Db     898 ACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTG 957

Qy     961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGC 1020
           ||||||||||||||||||||||||||||||||||||||||| || |||||||||||||||
Db     958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGCGCCAGC 1017

Qy    1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1018 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1077

Qy    1081 CTGGCCGAGGCGATGAGCCAGGCCAACACCAGCGTGATGATGCAGAAGAGCAACTTCAAG 1140
           ||||||||||| |||||||||||||||| || | |  || |||||   ||||||||||||
Db    1078 CTGGCCGAGGCCATGAGCCAGGCCAACAGCAACATCCTGGTGCAGCGCAGCAACTTCAAG 1137

Qy    1141 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200
           |||  |   |||||| |||||||||||||||||||||||| |||||||||||||||||||
Db    1138 GGCAGCAACCGCATCATCAAGTGCTTCAACTGCGGCAAGGTGGGCCACATCGCCCGCAAC 1197

Qy    1201 TGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
           ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    1198 TGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCCAGGAGGGCCACCAGATGAAG 1257

Qy    1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1258 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1317

Qy    1321 CCCGGCAACTTCCTGCAGAGCCGCC---------------------CCGAGCCCACCGCC 1359
           ||||||||||||||||||| |||||                     ||||||||||||||
Db    1318 CCCGGCAACTTCCTGCAGAACCGCCCCGAGCCCACCGCCCCCCCCGCCGAGCCCACCGCC 1377

Qy    1360 CCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGGCCAGAAGCAGGAGAGCAAG 1419
           ||||||||||||||||||||||||||||||||||||||||  |      |||||   |||
Db    1378 CCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGTGCCCCGCAAGGAG---AAG 1434

Qy    1420 GACCGCGAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA 1479
           || |||||| |||||||||||||||||||||||||||||| |||||||||||||||||||
Db    1435 GAGCGCGAGCCCCTGACCAGCCTGAAGAGCCTGTTCGGCAGCGACCCCCTGAGCCAGTAA 1494

RESULT 8
AX455888
LOCUS       AX455888                1509 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 4 from Patent WO0204493.
ACCESSION   AX455888
VERSION     AX455888.1  GI:21714881
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 4 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..1509
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="synthetic Gag of HIV strain AF110967"
ORIGIN

  Query Match             86.3%;  Score 1276.8;  DB 6;  Length 1509;
  Best Local Similarity   92.5%;  Pred. No. 2.9e-129;
  Matches 1399;  Conservative    0;  Mismatches   77;  Indels   36;  Gaps    4;

Qy       1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60
           |||||||||||||||||||||||||||||||  |||||||||   ||||||   ||||||
Db       1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGAGAAGCTGGACAAGTGGGAGAAGATCCGC 60
```

```
Qy      61 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
           |||||||||||||||||||||  ||||||| |||||||||||||||||||||||||||||
Db      61 CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120

Qy     121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
           ||||||   ||||||||||||||||||||||||||||||  |||||||||||||||||||
Db     121 CTGGAGGGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCAAGCAGATC 180

Qy     181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240
           ||    |||||||| ||||||||||||||||||| ||||||||||   ||||||| ||||
Db     181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC 240

Qy     241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300
           ||||||||||||||||||||||||||||     |||||||||||||||||||||||||||
Db     241 ACCGTGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTCCGCGACACCAAGGAGGCC 300

Qy     301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
           ||||||||||||||||||||||||||||||| ||||||||||| |||||||||| |||
Db     301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAG 360

Qy     361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
           ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GCCGAC---GGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 417

Qy     421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     418 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 477

Qy     481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     478 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 537

Qy     541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     538 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 597

Qy     601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCC 660
           |||||||||||||||||||||||||||||||||||| ||||||||||||| |||||||||
Db     598 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCAGGCCGGCCCC 657

Qy     661 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 720
            | |||||||||||||||||||| |||||||||||||||||||||||| |||||||||||
Db     658 GTGGCCCCCGGCCAGATGCGCGACCCCCGCGGCAGCGACATCGCCGGCGCCACCAGCACC 717

Qy     721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780
           ||||||||||||||||||||||||||||||||||||||| | ||||||||||||||||||
Db     718 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCGTGCCCGTGGGCGACATCTAC 777

Qy     781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     778 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 837

Qy     841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900
           |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db     838 CTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 897

Qy     901 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
           ||||||||||||||||||  ||||||||| |||||||||||||||||||| |||||||||
Db     898 ACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTG 957

Qy     961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGC 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCACC 1017

Qy    1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
           |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    1018 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1077

Qy    1081 CTGGCCGAGGCGATGAGCCAGGCCAACACC---AGCGTGATGATGCAGAAGAGCAACTTC 1137
           |||||||||||||||||||||||||||| |   | | | |||||||||||||||||||||
Db    1078 CTGGCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAGAAGAGCAACTTC 1137

Qy    1138 AAGGGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGC 1197
```

```
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      1138 AAGGGCCCCCGGCGCAACGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAG 1197

Qy      1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257

Qy      1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1317
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1317

Qy      1318 CGCCCCGGCAACTTCCTGCAGAGCCG------------------CCCCGAGCCCACCGCC 1359
             |||||||||||||||||||||| |||                  | ||| ||||||||||
Db      1318 CGCCCCGGCAACTTCCTGCAGAACCGCAGCGAGCCCGCCGCCCCCACCGTGCCCACCGCC 1377

Qy      1360 CCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGGCCAGAAGCAGGAGAGCAAG 1419
             |||||||||||||||||||||||||||||||||||||||| ||  |||||||||  ||||
Db      1378 CCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGCCCCCAAGCAGGAGCCCAAG 1437

Qy      1420 G------------ACCGCGAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCC 1467
             |            |||||||| ||||||||  ||||   ||||||||||||| || ||||
Db      1438 GACCGCGAGCCCTACCGCGAGCCCCTGACCGCCCTGCGCAGCCTGTTCGGCAGCGGCCCC 1497

Qy      1468 CTGAGCCAGTAA 1479
             ||||||||||||
Db      1498 CTGAGCCAGTAA 1509

RESULT 9
AX468544
LOCUS       AX468544                1509 bp    DNA     linear   PAT 16-JUL-2002
DEFINITION  Sequence 64 from Patent WO0226209.
ACCESSION   AX468544
VERSION     AX468544.1  GI:21901374
KEYWORDS    .
SOURCE      Human immunodeficiency virus 1 (HIV-1)
  ORGANISM  Human immunodeficiency virus 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1
  AUTHORS   O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
            Singh,M., Ulmer,J. and Dubensky,T.W.
  TITLE     Microparticles for delivery of the heterologous nucleic acids
  JOURNAL   Patent: WO 0226209-A 64 04-APR-2002;
            CHIRON CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1509
                     /organism="Human immunodeficiency virus 1"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11676"
ORIGIN

  Query Match             86.2%;  Score 1275.2;  DB 6;  Length 1509;
  Best Local Similarity   92.5%;  Pred. No. 4.3e-129;
  Matches 1398;  Conservative    0;  Mismatches   78;  Indels   36;  Gaps    4;

Qy         1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60
             |||||||||||||||||||||||||||||||  |||||||||   ||||||   ||||||
Db         1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGAGAAGCTGGACAAGTGGGAGAAGATCCGC 60

Qy        61 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
             |||||||||||||||||||||  ||||||| ||||||||||||||||||||||||| |||
Db        61 CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGTGAG 120

Qy       121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
             ||||||   ||||||||||||||||||||||||||||||  |||||||||||||||||||
Db       121 CTGGAGGGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCAAGCAGATC 180

Qy       181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240
             ||    |||||||| ||||||||||||||||||| ||||||||||   ||||||| ||||
Db       181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC 240

Qy       241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300
             ||||||||||||||||||||||||||||     |||||||||||||||||||||||||||
Db       241 ACCGTGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTCCGCGACACCAAGGAGGCC 300

Qy       301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
             ||||||||||||||||||||||||||||||| ||||||||||| |||||||||| |||
Db       301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAG 360

Qy       361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
             ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 GCCGAC---GGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 417

Qy       421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       418 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 477

Qy       481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       478 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 537

Qy       541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       538 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 597

Qy       601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCC 660
             |||||||||||||||||||||||||||||||||||| ||||||||||||| |||||||||
Db       598 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCAGGCCGGCCCC 657

Qy       661 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 720
              | |||||||||||||||||||| |||||||||||||||||||||||| |||||||||||
Db       658 GTGGCCCCCGGCCAGATGCGCGACCCCCGCGGCAGCGACATCGCCGGCGCCACCAGCACC 717

Qy       721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780
             ||||||||||||||||||||||||||||||||||||||| | ||||||||||||||||||
Db       718 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCGTGCCCGTGGGCGACATCTAC 777

Qy       781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       778 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 837

Qy       841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900
             |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db       838 CTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 897

Qy       901 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
             ||||||||||||||||||  ||||||||| |||||||||||||||||||| |||||||||
Db       898 ACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTG 957

Qy       961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGC 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       958 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCACC 1017

Qy      1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
             |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db      1018 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTG 1077

Qy      1081 CTGGCCGAGGCGATGAGCCAGGCCAACACC---AGCGTGATGATGCAGAAGAGCAACTTC 1137
             |||||||||||||||||||||||||||| |   | | | |||||||||||||||||||||
Db      1078 CTGGCCGAGGCGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAGAAGAGCAACTTC 1137

Qy      1138 AAGGGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGC 1197
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      1138 AAGGGCCCCCGGCGCAACGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAG 1197

Qy      1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257

Qy      1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1317
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1317

Qy      1318 CGCCCCGGCAACTTCCTGCAGAGCCG------------------CCCCGAGCCCACCGCC 1359
             |||||||||||||||||||||| |||                  | ||| ||||||||||
```

```
RESULT 10
LOCUS       AX468548                1509 bp    DNA     linear   PAT 16-JUL-2002
DEFINITION  Sequence 68 from Patent WO0226209.
ACCESSION   AX468548
VERSION     AX468548.1  GI:21901378
KEYWORDS
SOURCE      Human immunodeficiency virus 1 (HIV-1)
  ORGANISM  Human immunodeficiency virus 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1
  AUTHORS   O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
            Singh,M., Ulmer,J. and Dubensky,T.W.
  TITLE     Microparticles for delivery of the heterologous nucleic acids
  JOURNAL   Patent: WO 0226209-A 68 04-APR-2002;
            CHIRON CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1509
                     /organism="Human immunodeficiency virus 1"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11676"

ORIGIN

Query Match           86.1%;  Score 1273.6;  DB 6;  Length 1509;
Best Local Similarity 91.8%;  Pred. No. 6.4e-129;
Matches 1388;  Conservative 12;  Mismatches 76;  Indels 36;  Gaps 4;
```

```
Qy     1 ATGGGGGCCCGGGCCCAGCATCCTGCGCGGGGCCCGAAGCCCAGCGGGCATCCGC     60
         ||||||||||||||||||||||||||||||||||||||||||||  |||||||
Db  1018 ATGGGGGCCCGGGCCCCGCCAGCATCCTGCGGGGGCCCGAGCGGGCATCCGC   1077

Qy    61 CTGCGCGCCCCGGGGCGGCCAAGAAGCACTACATGCGTGTGTGGCCCAGCCCGCGAG    120
         |||||||||||||||||||||||||||||||||||||||||||
Db  1138 AAGGGCCCCCGAGGGCCATCGCCAAGTGCTTCAAACTGCCAAGCACATCGCCAAG   1197

Qy   121 CTGGGAGAAGTTCGCCCTGAACCCCGGGCCCTGCTCGGACGCCTGCAAGCAGATC    180
         |||||||||||||||||||||||||||||||||||||||||
Db  1198 AACTGCCGCGCCGGCCGCCCCGAAGGGCCTGCTGGGAATGTGGGCCAAGGGCCAAG   1257

Qy   181 ATCCGCGCCAGCTGCACCCCCGGCCCCTGCAGCCCGCCCGCGAGGAGCCTGTTCAAC    240
         ||||| |||||||||
Db  1258 AAGGACTGCACCGAGCGCCAGGCCGCCAACTTCCTGGGCAAGATCTGGCCCAGCC   1317

Qy   241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAGATCAGAGGTCCGCGACACCAAGGAGGCC    300
         ||||||||||||
Db  1318 CGCCCCGGGCCAACTTCCTGCGAGAACCGCAGCAGGTGCGCCACCGC   1377

Qy   301 CTGGACAAAGATCGAGGAGGACGAGAACAAGWCCAGCAGCAGGCCAAGGAG    360
         ||||||||||:::|||||
Db  1378 CCCCCGGGCCCGGAGAGCTTCGGCTTCGGCGAGGAGCAGGAGGGCCAAG   1437

Qy   361 GCCGAGAAGGGGGCCAAGGCTTGAGCGGAGAACTACCCCATCGTGCAGAACGATATG    420
         ||||||||||||||||||||||||||||||
Db  1438 GACCCGGCAGCCCCTACCCGCGAGCCCCTTGCGGCAGCCCTGCAGCGGCGCC   1497

Qy   421 GCCGGAC---GGGAAGGTGAGCCAGAACTACCCCCATCGTGAAGCTGGGGTGAAGGTGAGGGTGATCGAGGAGGAAG    480
         |||||||     ||||||||||||||||||||||||
Db  1498 CTGGAGCCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGAGGGTGATCGAGGAGAAG   1498
```

```
RESULT 11
LOCUS       AX455905              1509 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 21 from Patent WO0204493.
ACCESSION   AX455905
VERSION     AX455905.1  GI:21714897
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 21 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..1509
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="synthetic Gag coding sequence of HIV strain
                     AF110967"
ORIGIN

Query Match             85.2%;  Score 1260.8;  DB 6;  Length 1509;
Best Local Similarity   91.9%;  Pred. No. 1.5e-127;
Matches 1389; Conservative   0; Mismatches  87; Indels  36; Gaps   4;
```

```
Qy         1 ATGGGGCGCCCCGGCCAGCATCCTGCGCGGGGGCGCCAAGCTGGGACGCCATCCGC   60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGGGGCGCCCCGGCCAGCATCCTGCGCGGGGGCGCCAAGCTGGGACGCCATCCGC   60
```

```
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 168 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..9166
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

  Query Match             82.6%;  Score 1221.2;  DB 6;  Length 9166;
  Best Local Similarity   90.1%;  Pred. No. 1.7e-123;
  Matches 1346;  Conservative    0;  Mismatches  133;  Indels   15;  Gaps    3;

Qy          1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1880 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCAAG 1939

Qy         61 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
              |||   |||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db       1940 CTGAAGCCCGGCGGCAAGAAGCACTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 1999

Qy        121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
              ||||||   ||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       2000 CTGGAGCGCTTCGCCCTGGACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 2059

Qy        181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAAC 240
              ||    |||||||| ||||||||||||||||||| | |||||||||  ||||||  ||||
Db       2060 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCAAGGAGCTGATCAGCCTGCACAAC 2119

Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300
              |||||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db       2120 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGACGTGCGCGACACCAAGGAGGCC 2179

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
              |||||||||||| ||||||||||||||||| |||||||||||| ||||||||||||||||
Db       2180 CTGGACAAGATCAAGGAGGAGCAGAACAAGAGCCAGCAGAAGACCCAGCAGGCCGAGGCC 2239

Qy        361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2240 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 2299

Qy        421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
              ||||||||| |||||||| ||||||||||||||||||||||||||||| | |||||||||
Db       2300 GTGCACCAGCCCATCAGCGCCCGCACCCTGAACGCCTGGGTGAAGGTGGTGGAGGAGAAG 2359

Qy        481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2360 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 2419

Qy        541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600
              ||||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||
Db       2420 GACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 2479

Qy        601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCC 660
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       2480 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCACGCCGGCCCC 2539

Qy        661 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 720
               | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2540 GTGGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 2599

Qy        721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780
              ||||||||||||||| |||||||||||| |||||||||| | ||||||||||||||||||
Db       2600 CTGCAGGAGCAGATCACCTGGATGACCAACAACCCCCCCGTGCCCGTGGGCGACATCTAC 2659

Qy        781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
              ||||| |||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db       2660 AAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACAGCCCCGTGAGCATC 2719

Qy        841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2720 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 2779

Qy        901 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
              ||||||||||||||||  ||||||||| ||||||||||||||||||||||||||||||||
Db       2780 ACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTG 2839

Qy        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGC 1020
              ||||||||||||||||||||||||||||||||||||||||| || |||||||||||||||
Db       2840 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGCGCCAGC 2899

Qy       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       2900 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGCGGCCCCGGCCACAAGGCCCGCGTG 2959

Qy       1081 CTGGCCGAGGCGATGAGCCAGG---CCAACACCAGCGTGATGATGCAGAAGAGCAACTTC 1137
              ||||||||||| |||||| |||    |||||||| | | |||||||||   ||||||| |
Db       2960 CTGGCCGAGGCCATGAGCAAGGTGAACAACACCAACATCATGATGCAGCGCAGCAACTGC 3019

Qy       1138 AAGGGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGC 1197
              |||||||||  ||||| | |||||||||||||||||||||||||||||||| | ||||||
Db       3020 AAGGGCCCCAAGCGCACCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCTGGCCCGC 3079

Qy       1198 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATG 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       3080 AACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGGTG 3139

Qy       1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGC 1317
              ||||||||||||||||| ||||| || ||  |  | |||||||||||   |||||||||
Db       3140 AAGGACTGCACCGAGCGACAGGCTAATTTTTTTAGGAAGATCTGGCCTTCCCACAAGGGA 3199

Qy       1318 CGCCCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCCCCCGCCGAGAGCTTC 1377
               | || || || || || |||||| | || ||||| || ||||| || |  |||||||||
Db       3200 AGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAAGAGAGCTTC 3259

Qy       1378 CGCTTCGAGGA------GACCACCCCCGGCCAGAAGCAGGAG------AGCAAGGACCGC 1425
               | || | |||       || || |||   ||||||||||||        |||||| |
Db       3260 AGGTTTGGGGAAGAGACAACAACTCCCTCTCAGAAGCAGGAGCCGATAGACAAGGAACTG 3319

Qy       1426 GAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA 1479
               |  |  |  |   ||| |     || || |||| |||||||  |   || |||
Db       3320 TATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAA 3373

RESULT 13
AX427936
LOCUS       AX427936                9788 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 174 from Patent WO0232943.
ACCESSION   AX427936
VERSION     AX427936.1  GI:21538023
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 174 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..9788
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

  Query Match             81.5%;  Score 1206;  DB 6;  Length 9788;
  Best Local Similarity   90.0%;  Pred. No. 7.5e-122;
  Matches 1344;  Conservative    0;  Mismatches  130;  Indels   20;  Gaps    4;
```

```
Qy      1 ATGGGGCGCCCCGCGCCCAGCATCCTTGCGCGGCGGCCGCCAGCGTCGATGGAGCCGAAGACTTC    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1880 ATGGGGCGCCCCGCGCCCAGCATCCTTGCGCGGCGGCCGCCAGCGTCGATGGAGCGGCCATCAA  1939

Qy     61 CTGAAGCCCGGCGGCGGCGAAGAGTGCTACATGATGAAGAACGCACCTGGTGTGGGCCAGCGGGAG   120
          |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1940 CTGGACGGCTTCGCCCTGCCTGCTGGGGGCCTGCAAGAGGGCCTGCAAGCAGATC  1999

Qy    121 CTGGAGAAGTTCGGCCTGCCCTGAAACCCCGGGGCGGCCGCCAGCGTCGACAGATC   180
          ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2000 CTGGACGGCGCTTCGGCCCTGGACCCCGGGGCCTGCAAGCAGATC  2059

Qy    181 ATCCGGCCAGCGTGCACCCCGGCCCTGCAGAGGCGGCCAGCGGACCGCCAGGGCTGCAAGC   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2060 ATGAAGCAGCTGCACCCCGGCCCCTGCAGAGGCGGCCAGCGGACCGCCAGGGC  2119

Qy    241 ACCGGTGGCCCACCCGTACTGCGGTCACGGAGGAGGATCGAGGTCCGGCACGAGGAGGCGGCC   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2120 ACCGGTGGCCCACCCGTACTGCGGTGCACGAGGAGAGAAGCGGGCTTCAAG  2179

Qy    301 CTGGACAAGATCGAGGAGGAGGACGAGGACCAAGTGGCCAGCGGCCAGCGGCCGAGGGCC   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2180 CTGGACAAGATCAAGGAGGACGAGGACCAAGACCAGGCCGAGGGCC  2239

Qy    361 GCCGACAAGGGCCAAGGTGAGCCAGAAGCTACCCATCGTTGCAGTAAGGGCCCAGATG   420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2240 GCCGACAAGGGCCAAGGTGAGCCAGAACTACCCATCGTTGCAGTAAGGGCCCAGATG  2299

Qy    421 GTGCACCAGGCCATCAGCCCCGGACCTGGGGTCGAAGGTGATCGAGGAGAAG   480
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   2300 GTGCACCAGCCCATCAGCGCCCACCCTGAACGCTCCGGTGAAGGTGGTGGAAGGAAG  2359

Qy    481 GCCTTCAGCGCCCGGCCCCCGCCCATGTTCACCGCGCCCTGAGCGAGGAGGAGGCCC   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2360 GCCTTCAGCGCCCGGCCCCCCGCCCATGTTCACCGCCCGGCACCCCCAG  2419

Qy    541 GACCTGAACACGATAGTTGAACCCGTGGGCGGCCGCCACCCAGGCCCCAGTGCTTGAAG   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2420 GACTCGAACACCATGCTGACACGTGGCGCGCCACCGTGGACCCCATGCTCAGAGGGC  2479

Qy    601 GACACCATCAAGGAGGAGGCCGAGTGGGACGGCGTGCACCCCGTGCACCGGCGGCCCC   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2480 GACACCATCAACGAGGAGGCCGCCGAGTGGGACGGCCGCTGCACCCCGTGCACCGGCGGCCCC  2539

Qy    661 ATCGGCCCCCCCGGCCGCGATCGCCCGGAGCCCCGCGGGCGCGCACATCGCCACCAGCACC   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2540 GTGGGCCCCCCGGCAGTGCCGCGAGCCCCGGCCAGCCGCACCATCGCCACCAGCACC  2599

Qy    721 CTGCAGGAGCAGATCTCCCTGGATGACAGCAACCAACCCCCATCCCGTGGGCGGCACATCTAC   780
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2600 CTGCAGGAGCAGATCAGCCTGGATGACAGCAACCAACCCCCATCCGTGGGCGGACATCTAC  2659

Qy    781 AAGCGGTGGATCATCCTGGGCGGCTGAACAAGATCGTTGCGGGATGTACAGCAGCGTGAGCGATC   840
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2660 AAGGCCTGGATCATCCTGGGCGGCTGAACAAGATCGTTGCGGGTGTACAGCCGCCATGGCGATC  2719

Qy    841 CTGGACACATCAAGCACGAGGGCCCCCAAGAGCGCCCTTCCCGGACTTACGTGCGACCGCCTTCTTCAAG   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2720 CTGGACATCAAGCACGAGGGCCCCAAGAGCGCCCTTCCCGGACTTCCTGGACAAGATCAGG  2779

Qy    901 ACCCTTCGCGCCGAGCGAGGAGTTGAACAACTGGATCGACACCCTGCTG   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2780 ACCCTGCGCGCCGGACCCAGAGGCCACCCAGAGCGTGAAGAACTGGGATCGACACCCTGCTG  2839

Qy    961 GTGCAGAACGCCAACCCAGACCATCTCGGCGCTCGGCGGGCC   1020
          ||| ||||||||||||||||||||||||||||||||||||||||||
Db   2840 GTGGACAGAACGCCAACCCCGACTGCGAAGAGGGCCCTGCAAGCAGATC  2899

Qy   1021 CTGGAGGAGGAGATGATGACCGCGCTGCCAGGGGCTGGGGCGGGGGGGCCCCGCGTG   1080
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2900 CTGGAGGAGGAGGATGACCGCGCCGCCGAGGAGGCCTGCAAGCAGATC  2959
```

```
Qy   1081 CTGGCCGAGGCGGATGAGCCAGG---CCAACACCAGCCGGTGATGATGGAGAAGAGCAAACTTC   1137
          |||||||||||||||||||||||    |||||||
Db   2960 CTGGCCGAGGCCATGAGCAGCAGTGAACAGCACCAACATCATGATGGCAGCCAGCAACTGC  3019

Qy   1138 AAGGGCGCCCCCGGAGCGCCCAGCCAGCCATCGTCAAGGTGAACAAGGTGAACCACACACCAAC  1197
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3020 AAGGGCGCCCCAAGCACCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCTGGCCCGC  3079

Qy   1198 AACTGCCCGCGCCCCCCCCCGCAAGAAAGAGGCCTGGAAGTGGCGGCAAGGAGGCCAGCGATG  1257
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3080 AACTGCCGCGCGCCCCCCGCAAGAGGCCTGCTGGAAGTGGGGCAAGGAGGGGTG  3139

Qy   1258 AAGGACTGCACCCAGCGCCAGCGCCAGCCCAACTTCCTGCGCCAAGTTCCTGGGCCAGATCCTGCCCAGAGGGC  1317
          ||||||||||||||||||||||||||||||||||||  |||||||||||||||||||
Db   3140 AAGGACTGCACCCAGCGCCAGCGACCAGGCCTGA------TAGGGAGATCTCCCTGCCCCACAAGGGA  3194

Qy   1318 CGCCCCGGCCAACTTCCTGCGCAGCGCCCCGGCGAGCCCACCGCCGCCCCGCCGAGCTTC  1377
          ||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db   3195 AGGCCCAGGAATTTCTTCAGCGCGAGCGACCAGGCCACCAGCCCACCGCCCAGGAGCTTC  3254

Qy   1378 CGCTTCGACGA--------GACCACCCCCGGCCAGGAAGCACAGGAG   1425
          ||||||||||        |||||||||||||||||||||||||||
Db   3255 AGGTTTGGGGAGAACAACAACTCCCCTCAGAAGACCAAGAGCCGGATAGAGAAGGAAACTG  3314

Qy   1426 GAGACCCTGACCAGCCTGAAGAGACCCTGTTCGGCAACGACGACCCCCTGAGCCAGTAA  1479
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3315 TATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCCTCGTCACTATAA  3368
```

RESULT 14

```
LOCUS        AX149648            4288 bp    DNA      linear   PAT 08-JUN-2001
DEFINITION   Sequence 2 from Patent WO0136614.
ACCESSION    AX149648
VERSION      AX149648.1  GI:14348047
KEYWORDS
SOURCE       Human immunodeficiency virus
  ORGANISM   Human immunodeficiency virus
             Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
             lentivirus group.
REFERENCE    1
  AUTHORS    Shao,Y., Wagner,R., Wolf,H. and Graf,M.
  TITLE      The genome of the hiv-1 inter-subtype (c/b') and use thereof
  JOURNAL    Patent: WO 0136614-A 2 25-MAY-2001;
             Geneart GMBH Gesellschaft fuer angewandte Boitechnologie (DE) ;
             Shao, Yiming (CN)
FEATURES             Location/Qualifiers
     source          1..4288
                     /organism="Human immunodeficiency virus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:12721"
ORIGIN
```

```
Query Match         81.5%;  Score 1205.4;  DB 6;  Length 4288;
Best Local Similarity  89.5%;  Pred. No. 1.1e-121;
Matches 1335;  Conservative  0;  Mismatches 141;  Indels 15;  Gaps  3;
```

```
Qy      1 ATGGGGCGCCCCGCGCCCAGCATCCTTGCGCGGCGGCCGCCAGCGTCGATGGAGCCGAAG    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     13 ATGGGGCGCCCCGCGCCCAGCATCCTTGCGCGGCGGCCGCCAGCGTCGGAGATCAGG    72

Qy     61 CTGGCGCCCCGGCGGCGGCGAAGAGTGCTACATGATGAAGAACGCACCTGGTGTGGGCGCGGAG   120
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     73 CTGAGGCGCCCGGCGGCAAGGAGGCCACCTACATGATCAAGAAGAACACCTGTGTGTGGGCCGCGCTG   132

Qy    121 CTGGAGAAGTTCGGCCCTGAACCCCGGGCCGCCTGCACATGCTGAGAAGCACCTGGTGTGGTGGGGGCCGCGCTG    180
          || ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    133 CTGGAGGAGGTTCGCCCCTGAAACCCCGGGCCTGCAGAGAGGACCAGGCCGCCTGTTCAAC   192

Qy    181 ATCCGGCCAGCTGCACCCCCGGCCCCTGCAGACCGGCGGCCAGCGGACCCTGCAGAGGCTGTTCAAC   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    193 ATGAAGCAGCTGCAGAGGCGGCCTGCAGACCGGCACCGGCACGAGGAGCTGCAGAGGATC   252
```

```

```
Qy        241 ACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCC 300
              |||||||||||||  ||||||||||||    ||||||| ||  | ||||||| |||||||
Db        253 ACCGTGGCCACCCCCTACTGCGTGCACACCGAGATCGACGTGAGGGACACCAGGGAGGCC 312

Qy        301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCC 360
              ||||||||||||||||||||||||||||||  ||||||||||| |||||||||| |||  
Db        313 CTGGACAAGATCGAGGAGGAGCAGAACAAGATCCAGCAGAAGACCCAGCAGGCCAAGGAG 372

Qy        361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 GCCGAC---GGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG 429

Qy        421 GTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAG 480
              ||||||||| ||||||||||| | |||||||| || |||||||||||| | |||||||||
Db        430 GTGCACCAGCCCATCAGCCCCAGGACCCTGAATGCATGGGTGAAGGTGGTGGAGGAGAAG 489

Qy        481 GCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAG 540
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        490 GCCTTCAGCCCCGAGGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGCGCCACCCCCCAG 549

Qy        541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAG 600
              ||||||||||| ||| ||||||||||||||||||||||||||||||||||||| ||||||
Db        550 GACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATCCTGAAG 609

Qy        601 GACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCC 660
              ||||||||||||||||||||||||||||||||| |  |||||||||||||||||||||||
Db        610 GACACCATCAACGAGGAGGCCGCCGAGTGGGACAGGCTGCACCCCGTGCACGCCGGCCCC 669

Qy        661 ATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACC 720
              |||||||||||||||||| | |||||| | |||||||||||||||||||||||||||| |
Db        670 ATCGCCCCCGGCCAGATGAGGGAGCCCAGGGGCAGCGACATCGCCGGCACCACCAGCAAC 729

Qy        721 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTAC 780
              ||||||||||||||||||||||||||||||||||||||| | ||||||||||||||||||
Db        730 CTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCGTGCCCGTGGGCGACATCTAC 789

Qy        781 AAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
              ||| ||||||||||||||||  | |||||||||||| ||||||||||||||   ||||||
Db        790 AAGAGGTGGATCATCCTGGGTTTAAACAAGATCGTGAGGATGTACAGCCCCACCAGCATC 849

Qy        841 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAG 900
              ||||||||||||||||||||||||||||||||| | |||||||| ||| | |||||||||
Db        850 CTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCAGGGACTACGTCGACAGGTTCTTCAAG 909

Qy        901 ACCCTGCGCGCCGAGCAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTG 960
              |||||| | || ||||||  ||||||||  ||||||||||||||||||||||||||||||
Db        910 ACCCTGAGGGCGGAGCAGGCCACCCAGGGCGTGAAGAACTGGATGACCGACACCCTGCTG 969

Qy        961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGC 1020
              |||||||||||||||||||||||||||||||||||| | || || |||||||||||||||
Db        970 GTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAGGGCCCTGGGCCCCGGCGCCAGC 1029

Qy       1021 CTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
               | |||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db       1030 ATCGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCAAGGTG 1089

Qy       1081 CTGGCCGAGGCGATGAGCCAGGCCAACACCAGCGTGATGATGCAGAAGAGCAACTTCAAG 1140
              ||||||||||| ||||||||| |||||| |  | |  ||||||||| |||||||||||||
Db       1090 CTGGCCGAGGCCATGAGCCAGACCAACAGCGCCATCCTGATGCAGAGGAGCAACTTCAAG 1149

Qy       1141 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAAC 1200
              |||  |  | | ||||| |||||||||||||||||||||||||||||||||||| | |||
Db       1150 GGCAGCAAGAGGATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAGGAAC 1209

Qy       1201 TGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1260
              ||| | |||||| | |||||||||||||||||||||||||||||||||||||||||||||
Db       1210 TGCAGGGCCCCCAGGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAG 1269

Qy       1261 GACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGC 1320
              |||||||||||| | |||||||||||||||||||||||||||||||||||||||||| ||
Db       1270 GACTGCACCGAGAGGCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCGGC 1329

Qy       1321 CCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCCCCCGCCGAGAGCTTCCGC 1380
              ||||||||||||||||||| | | ||||||||||||||||||||||  ||||||||| | 
Db       1330 CCCGGCAACTTCCTGCAGAACAGGCCCGAGCCCACCGCCCCCCCCGAGGAGAGCTTCAGG 1389

Qy       1381 TTCGAGGAG------ACCACCCCCGGCCAGAAGCAGGAG------AGCAAGGACCGCGAG 1428
              |||||||||      ||||||||| ||||||||||||||        |||||| |   | 
Db       1390 TTCGAGGAGGAGACCACCACCCCCAGCCAGAAGCAGGAGCCCATCGACAAGGAGCTGTAC 1449

Qy       1429 ACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA 1479
               ||||||||||||||||||||||||||||||||||||||   |||||| ||
Db       1450 CCCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCCAGCAGCCAGGAA 1500

RESULT 15
AF201927
LOCUS       AF201927                1509 bp    DNA     linear   SYN 16-MAR-2000
DEFINITION  Synthetic construct gag protein gene, complete cds.
ACCESSION   AF201927
VERSION     AF201927.1  GI:7248702
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 1509)
  AUTHORS   zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,
            Selby,M., Otten,G.R. and Barnett,S.W.
  TITLE     Increased expression and immunogenicity of sequence-modified human
            immunodeficiency virus type 1 gag gene
  JOURNAL   J. Virol. 74 (6), 2628-2635 (2000)
  MEDLINE   20148954
   PUBMED   10684277
REFERENCE   2  (bases 1 to 1509)
  AUTHORS   zur Megede,J. and Barnett,S.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton,
            Emeryville, CA 94608, USA
FEATURES             Location/Qualifiers
     source          1..1509
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     CDS             1..1509
                     /note="HIV-1SF2 p55gag"
                     /codon_start=1
                     /transl_table=11
                     /product="gag protein"
                     /protein_id="AAF43628.1"
                     /db_xref="GI:7248703"
                     /translation="MGARASVLSGGELDKWEKIRLRPGGKKKYKLKHIVWASRELERF
                     AVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIDVKDTKEALE
                     KIEEEQNKSKKKAQQAAAAAGTGNSSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKV
                     VEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVH
                     PVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIV
                     RMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNANPDCKT
                     ILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNPATIMMQRGNFRNQRKTV
                     KCFNCGKEGHTARNCRAPRKKGCWRCGREGHQMKDCTERQANFLGKIWPSYKGRPGNF
                     LQSRPEPTAPPEESFRFGEEKTTPSQKQEPIDKELYPLTSLRSLFGNDPSSQ"
ORIGIN

  Query Match             81.0%;  Score 1197.8;  DB 12;  Length 1509;
  Best Local Similarity   89.3%;  Pred. No. 9.9e-121;
  Matches 1347;  Conservative    0;  Mismatches  132;  Indels   30;  Gaps    4;

Qy          1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGC 60
              |||||||||||||||||| | ||| |||||||| ||||||||   ||||||   ||||||
Db          1 ATGGGCGCCCGCGCCAGCGTGCTGAGCGGCGGCGAGCTGGACAAGTGGGAGAAGATCCGC 60

Qy         61 CTGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
              |||||||||||||||||||||   |||| | |||||||| | ||||||||||||||||||
Db         61 CTGCGCCCCGGCGGCAAGAAGAAGTACAAGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 120

Qy        121 CTGGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC 180
              ||||||   |||||| |||||||||||||||||||||||||||||||||||   ||||||
Db        121 CTGGAGCGCTTCGCCGTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCCGCCAGATC 180
```

Qy    181  ATCCGCCAGCTGCACCCGCCCTTGCAGAGGCTTGAAGAGCCTTGTTCAAC  240
           ||||||||||||||    |||||  |||||||||  ||  |||||||||| ||||
Db    181  CTGGGCCAGCTGCAGCCGCAGCTGCCGGAGGCTGCGCAGAGCCTTGTACAAC  240

Qy    241  ACCGTGGCCACCCTGTACTGCGTGCACGAGAGTCGAGGTCCGCGACACCAGGCC  300
           ||||||||||  ||||||||  |||||  |||||  ||||||||||||||||||
Db    241  ACCGTGGCCACCCTGTACTGCGTGCACGAGACGTCAAGGTACCGGAGGCC  300

Qy    301  CTGGACAAGGATCGAGGAGGAGGAACAAGTCCAGCAGACCAACAGGAGGCC  360
           |||  ||||||||||||||||||||||||||||||  |||||||  |||||||
Db    301  CTGGAGAAGGATCGAGGAGGAGCAGAACAAGTCCAGAGAGACCGGCGC  360

Qy    361  GCCG--------ACAAGGGCAAGGTGAGCCAAGACTACCCATGGTGCAGAAACTG  408
               |||||||||||||||||  |||||||||||||||||||||||||||
Db    361  GCCGCCGGCAACAGCGACCCAGGCGTGAGCCAAGACTACCCATGGTGCAGAAACTG  420

Qy    409  CAGGGCCCAGATGGTGCACCAGGCCATCAGCCCCCGGCCTGAACGCCAGGTG  468
           ||||||||||||||||||||||||||||||  |||||||||||||  |||||||
Db    421  CAGGGCCCAGATGGTGCACCCAGGCCATCAGCCCCCGGCCTGAACGCCCATGTG  480

Qy    469  ATCGAGGAGGAAGGCCTTCAGCCCCCGGAGGTGATCCCCATGTTCACCGGCCCTGAGCGAGGGC  528
           |||||||||||||||||  |||||||||||||||||||||||||||||||||||||||||
Db    481  GTCGGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCAGGGCCCTGAGCGAGGGC  540

Qy    529  GCCACCCCCCAGGACCTGAACACAGAGTTGAACACCGTGGGGCGGGGGCCACCAGGCCGCCATG  588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GCCACCCCCCAGGACCTGAACACAGAGTTCAGGGCCCTGAGGGGCGGGGGCCACCAGGCC  600

Qy    589  CAGATGCTGAAGGACACCATCAACGAGGAGGACCGCGGCGTGCACCCCGGTG  648
           ||||  |||||||||||||||||||||||||||||||||||||||||||||
Db    601  CAGATGCTGAAGGACCATCAACGAGGAGGACCGCGGCGTGCACCCCGGTG  660

Qy    649  CACGCCCGGCCCCCATCGCCCCCGGCCCCAGTGCCGGACAGCGATTCGCCGGC  708
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  CAGGCCGGCCCCATCGCCCCCAGATGCCGGCCAGCGACATCGCCGGC  720

Qy    709  ACCACCAGCACCCTGCAGGAGGACGAATCGCCCTGAACAAGCGCCCCCATCCCCGTG  768
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  ACCACCAGCACCCTGCAGGAGCAGATCGGCTGAATGACCGCCCCCCATCCCCGTG  780

Qy    769  GGCGCACATCTACAAGCCGGATCATCCTGGGGCCTGAACAAGATCGTCCGGATGTACAGC  828
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  GGCGAGATCTACAAGCCGGTGGATCATCCTGGGCCTGAACAAGATCGTCCGGATGTACAGC  840

Qy    829  CCCGTGAGCGCATCCTGGACATCAAGCAGGGGCCCCCAAGAGCCCTTCCGGCCGACTACGTGGAC  888
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  CCACCAGCATCCTGGACATCCCCCAGGGACCCCTTCCGGCCGACTACGTGGAC  900

Qy    889  CGCTTCTTCAAGACCCTTCGCGCGCGGAGCGAGGAGCACCCAGGAGGAGTGAAGAACTGGATGACC  948
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  CGCTTCTACAAGACCCTGCGTGCGCAGCCAGGCCAGGAGCTGAGAAGAACTGGATGACC  960

Qy    949  GACACCCTGCTGTGCAGAACCCCAACCCCGACTGCAAGAGACCATCCTGCGCGCTCTCGGC  1008
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  GAGACCCTGCTGTGCAGAAACGCCAACCCCGACTGCAAGAGACCATCCTGAAAGGCTTCTCGGC  1020

Qy    1009 CCCGGGGGCCCAGCCTGGAGGAGGCAGTGATGACCGGCTGCCGGGGCTGGCCGCGCCCCCAGCCAC  1068
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CCCGCGCCACCCTGGAGGAGGATGATGACCGCTGCCGGGGCTGGCGGCGCCCCAC  1080

Qy    1069 AAGGCCCCGTGCTGGGCGGAGGCCATGAGCGCAGGCCGAACA------CCAGGCGTGATGATG  1122
           ||||||||||||||||||||||||||||||||||||  ||||||||||||
Db    1081 AAGGCCGCGGTGCTGGGCGGGATGAGCCAGGTGACGAACCGGCACCATCATGAT  1140

Qy    1123 CAGAAGAGCAACTTCAAGGAGGCCCCGGGCCATCGTCAAGTGCTTCAACTCACTCGGGCAAG  1182
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CAGGGCGCGGCAACTTCCCGCAACCGGTCAAGTGCTTCAACTCGCGCGAAGTCGGGCAAGGAG  1200

Qy    1183 GGCCACATCGCCGCCCACCTGCCGCGGCCCCCCAAGAAGGGGCTGCTGCGAAGTTCGGCGCCGC  1242
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GGCCACACCGGCCCAACTGCCGCCCCCCGCAAGAAGGGGCTGCTGCGAAGGGCGCCGC  1260
```

```
Qy    1243 GAGGGCCACCAAGGATGAAGGACTGCACCCGAGGCGCCAGGCCAAGATCTGG  1302
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GAGGGCCACCAGATGAAGGACTGCACCGACGGCGCCAGGCCAAGATCTGG  1320

Qy    1303 CCCAGCCACAAGGGCGCCCGGCCCCGCCAACTTCCTGCAGAGGCCCCGGCCCC  1362
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CCCAGCTACAAGGGCGCGCCCCGGGGCCAACTTCCTGCAGAGGCCCACCGGCCCC  1380

Qy    1363 CCCGGCCGAGAGCTTCCGCTTC------GAGGAGCACCCCCGGCCGAAGGCAGGAG---  1413
           ||||  |||||||||||||||||    |||||||||||||||||||||  |||
Db    1381 CCCGGAGGAGCTTCCGCTTCGGGCAGGACACCCCCGGCCGAGAAGGCAGGAGGCCC  1440

Qy    1414 ---AGCAAGGACCCCGGAGACCCTGACCAGCCTGAAGGCTGTGCCGCAGACCACCCCCTG  1470
               |||||||||||||||||||||||||||||||||||||||||  ||||||||
Db    1441 ATCGACAAGGAGCTGTACCCCCTGACCAGCCTTGTTCGGCCAACGACCCCCAGC  1500

Qy    1471 AGCCAGTAA  1479
           |||||||||
Db    1501 AGCCAGTAA  1509
```

Search completed: May 28, 2004, 13:23:24
Job time : 6085.87 secs

```
QM nucleic - nucleic search, using sw model

                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:      May 28, 2004, 09:24:06 ; Search time 6.49156 Seconds
                                       (without alignments)
                                       5129.286 Million cell updates/sec

Title:          US-09-475-704A-2
Perfect score:  60
Sequence:       1 gacatccgccagggcccaa.........tggaccgcttcttcaagacc 60
Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:    1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued Patents NA:*
        1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
        2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
        3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
        4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
        5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
        6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result          Query
No.    Score    Match  Length  DB  ID                 Description
-----  -----    -----  ------  --  --                 -----------
  1    58.4     97.3     60    4   US-09-475-515-20   Sequence 20, Appl
  2    58.4     97.3   1268    4   US-09-475-515-9    Sequence 9, Appli
  3    58.4     97.3   1515    4   US-09-475-515-4    Sequence 4, Appli
  4    58.4     97.3   1853    4   US-09-475-515-5    Sequence 5, Appli
  5    58.4     97.3   1865    4   US-09-475-515-78   Sequence 78, Appl
  6    58.4     97.3   1865    4   US-09-475-515-79   Sequence 79, Appl
  7    58.4     97.3   2031    4   US-09-475-515-7    Sequence 7, Appli
  8    58.4     97.3   4319    4   US-09-475-515-6    Sequence 6, Appli
  9    58.4     97.3   4472    4   US-09-475-515-75   Sequence 75, Appl
 10    58.4     97.3   4608    4   US-09-475-515-76   Sequence 76, Appl
 11    58.4     97.3   4689    4   US-09-475-515-74   Sequence 74, Appl
 12    58.4     97.3   4766    4   US-09-475-515-73   Sequence 73, Appl
 13    47.8     79.7   4307    4   US-09-552-950-2    Sequence 2, Appli
 14    47.8     79.7   9772    4   US-09-552-950-5    Sequence 5, Appli
 15    58.3     98.3   1476    4   US-09-184-418C-100 Sequence 100, App
 16    58.3     98.3   1485    4   US-09-184-418C-73  Sequence 73, Appl
 17    35       58.3   8959    4   US-09-184-418C-11  Sequence 11, Appl
 18    35       58.3   9010    4   US-09-184-418C-8   Sequence 8, Appl
 19    33.4     55.7   718     6   5204259-2          Patent No. 5204259
 20    33.4     55.7   793     6   5204259-4          Patent No. 5204259
 21    33.4     55.7   796     6   5204259-6          Patent No. 5204259
 22    33.4     55.7   845     1   US-08-589-446-3    Sequence 3, Appli
 23    33.4     55.7   845     1   US-08-444-882-3    Sequence 3, Appli
 24    33.4     55.7   845     3   US-08-389-459A-3   Sequence 3, Appli
 25    33.4     55.7   845     3   US-08-087-867A-3   Sequence 3, Appli
 26    33.4     55.7   871     6   5204259-10         Patent No. 5204259
 27    33.4     55.7   1021    6   5204259-8          Patent No. 5204259
 28    33.4     55.7   4307    4   US-09-552-950-1    Sequence 1, Appli
 29    33.4     55.7   4338    4   US-09-872-733A-1   Sequence 1, Appli
 30    33.4     55.7   5362    3   US-08-463-210-5    Sequence 5, Appli
 31    33.4     55.7   5362    4   US-08-463-028-5    Sequence 5, Appli
 32    33.4     55.7   7228    2   US-08-850-049-128  Sequence 128, App
 33    33.4     55.7   7228    2   US-08-850-049-129  Sequence 129, App
 34    33.4     55.7   7228    2   US-08-050-478-128  Sequence 128, App
 35    33.4     55.7   7228    3   US-08-050-478-129  Sequence 129, App
 36    33.4     55.7   7228    3   US-09-414-117-128  Sequence 128, App
 37    33.4     55.7   7228    4   US-09-414-117-129  Sequence 129, App
 38    33.4     55.7   8366    2   US-09-678-437-128  Sequence 128, App
 39    33.4     55.7   8366    4   US-09-678-437-129  Sequence 129, App
 40    33.4     55.7   7399    2   US-09-872-733A-6   Sequence 6, Appli
 41    33.4     55.7   7228    4   US-08-188-583-5    Sequence 5, Appli
 42    33.4     55.7   9709    2   US-08-418-848A-9   Sequence 9, Appli
 43    33.4     55.7   9709    3   US-08-388-353-1    Sequence 1, Appli
 44    33.4     55.7   9709    4   US-08-488-551B-1   Sequence 1, Appli
 45    33.4     55.7   9709    4   US-09-309-572-15   Sequence 15, Appl
```

ALIGNMENTS

```
RESULT 1
US-09-475-515-20

;  Sequence 20, Application US/09475515A
;  Patent No. 6602705
;  GENERAL INFORMATION:
;  APPLICANT: BARNETT, Susan
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: SRIVASTAVA, Indresh
;  APPLICANT: LIAN, Ying
;  APPLICANT: HARTOG, Karin
;  APPLICANT: LIU, Hong
;  APPLICANT: GREER, Catherine
;  APPLICANT: SELBY, Mark
;  APPLICANT: WALKER, Christopher
;  TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;  TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;  FILE REFERENCE: 1621.002
;  CURRENT APPLICATION NUMBER: US/09/475,515A
;  CURRENT FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 90
;  SOFTWARE: PatentIn Ver. 2.0

;  SEQ ID NO 20
;  LENGTH: 60
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: synthetic p55
;  OTHER INFORMATION: Gag Major Homology Region
US-09-475-515-20

Query Match             97.3%;  Score 58.4;  DB 4;  Length 60;
Best Local Similarity   98.3%;  Pred. No. 2.7e-10;
Matches  59;  Conservative  0;  Mismatches 1;  Indels  0;  Gaps  0;

Qy    1 GACATCCGCCAGGGCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60


RESULT 2
US-09-475-515-9

;  Sequence 9, Application US/09475515A
;  Patent No. 6602705
;  GENERAL INFORMATION:
;  APPLICANT: BARNETT, Susan
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: SRIVASTAVA, Indresh
;  APPLICANT: LIAN, Ying
;  APPLICANT: HARTOG, Karin
;  APPLICANT: LIU, Hong
```

APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: Patentin Ver. 2.0
  SEQ ID NO 9
  LENGTH: 1268
  TYPE: DNA
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
  OTHER INFORMATION: common region
US-09-475-515-9

Qy   Query Match   97.3%;   Score 58.4;   DB 4;   Length 1268;
     Best Local Similarity   98.3%;   Pred No 3.5e-10;
     Matches   59;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Db
   1 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCCGACTAGGTGGACCGGCTTCTTCAAGACC   60
     ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
 862 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCGGACTAGGTGGACCGGCTTCTTCAAGACC  921

RESULT 3
US-09-475-515-4
Sequence 4, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: Patentin Ver. 2.0
  SEQ ID NO 4
  LENGTH: 1515
  TYPE: DNA
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-475-515-4

Qy   Query Match   97.3%;   Score 58.4;   DB 4;   Length 1515;
     Best Local Similarity   98.3%;   Pred. No 3.6e-10;
     Matches   59;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Db
   1 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCCGACTAGGTGGACCGGCTTCTTCAAGACC   60
     ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
 862 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCGGACTAGGTGGACCGGCTTCTTCAAGACC  921

RESULT 4
US-09-475-515-5
Sequence 5, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan

---

APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: Patentin Ver. 2.0
  SEQ ID NO 5
  LENGTH: 1853
  TYPE: DNA
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: synthetic
  OTHER INFORMATION: HIV-Gag-protease
US-09-475-515-5

Qy   Query Match   97.3%;   Score 58.4;   DB 4;   Length 1853;
     Best Local Similarity   98.3%;   Pred. No 3.7e-10;
     Matches   59;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Db
   1 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCCGACTAGGTGGACCGGCTTCTTCAAGACC   60
     ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
 862 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCGGACTAGGTGGACCGGCTTCTTCAAGACC  921

RESULT 5
US-09-475-515-78
Sequence 78, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: Patentin Ver. 2.0
  SEQ ID NO 78
  LENGTH: 1865
  TYPE: DNA
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: GP1
US-09-475-515-78

Qy   Query Match   97.3%;   Score 58.4;   DB 4;   Length 1865;
     Best Local Similarity   98.3%;   Pred. No 3.7e-10;
     Matches   59;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Db
   1 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCCGACTAGGTGGACCGGCTTCTTCAAGACC   60
     ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
 868 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCGGACTAGGTGGACCGGCTTCTTCAAGACC  927

RESULT 6
US-09-475-515-79

```
;  Sequence 79, Application US/09475515A
;  Patent No. 6602705
;  GENERAL INFORMATION:
;      APPLICANT: BARNETT, Susan
;      APPLICANT: ZUR MEGEDE, Jan
;      APPLICANT: SRIVASTAVA, Indresh
;      APPLICANT: LIAN, Ying
;      APPLICANT: HARTOG, Karin
;      APPLICANT: LIU, Hong
;      APPLICANT: GREER, Catherine
;      APPLICANT: SELBY, Mark
;      APPLICANT: WALKER, Christopher
;      TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;      TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;      FILE REFERENCE: 1621.002
;      CURRENT APPLICATION NUMBER: US/09/475,515A
;      CURRENT FILING DATE: 1999-12-30
;      NUMBER OF SEQ ID NOS: 90
;      SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 79
;      LENGTH: 1865
;      TYPE: DNA
;      ORGANISM: Artificial Sequence
;      FEATURE:
;      OTHER INFORMATION: Description of Artificial Sequence: GP2
US-09-475-515-79

  Query Match             97.3%;  Score 58.4;  DB 4;  Length 1865;
  Best Local Similarity   98.3%;  Pred. No. 3.7e-10;
  Matches   59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy          1 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCGACTAGGTGGACCGGCTTCTTCAAGACC 60
              |||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||
Db        868 GACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 927


RESULT 7
US-09-475-515-7
;  Sequence 7, Application US/09475515A
;  Patent No. 6602705
;  GENERAL INFORMATION:
;      APPLICANT: BARNETT, Susan
;      APPLICANT: ZUR MEGEDE, Jan
;      APPLICANT: SRIVASTAVA, Indresh
;      APPLICANT: LIAN, Ying
;      APPLICANT: HARTOG, Karin
;      APPLICANT: LIU, Hong
;      APPLICANT: GREER, Catherine
;      APPLICANT: SELBY, Mark
;      APPLICANT: WALKER, Christopher
;      TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;      TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;      FILE REFERENCE: 1621.002
;      CURRENT APPLICATION NUMBER: US/09/475,515A
;      CURRENT FILING DATE: 1999-12-30
;      NUMBER OF SEQ ID NOS: 90
;      SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 7
;      LENGTH: 2031
;      TYPE: DNA
;      ORGANISM: Artificial Sequence
;      FEATURE:
;      OTHER INFORMATION: Description of Artificial Sequence: synthetic
;      OTHER INFORMATION: HIV-Gag/HCV-core fusion polypeptide
US-09-475-515-7

  Query Match             98.3%;  Score 58.4;  DB 4;  Length 2031;
  Best Local Similarity   98.3%;  Pred. No. 3.7e-10;
  Matches   59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy          1 GACATCCGCCAGGGGCCCAAGGAGCCCTTCCGCGACTAGGTGGACCGGCTTCTTCAAGACC 60
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        862 GACATCCGCCAGGGGCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCTACAAGACC 921
```

```
RESULT 8
US-09-475-515-6
;  Sequence 6, Application US/09475515A
;  Patent No. 6602705
;  GENERAL INFORMATION:
;      APPLICANT: BARNETT, Susan
;      APPLICANT: ZUR MEGEDE, Jan
;      APPLICANT: SRIVASTAVA, Indresh
;      APPLICANT: LIAN, Ying
;      APPLICANT: HARTOG, Karin
;      APPLICANT: LIU, Hong
;      APPLICANT: GREER, Catherine
;      APPLICANT: SELBY, Mark
;      APPLICANT: WALKER, Christopher
;      TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;      TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;      FILE REFERENCE: 1621.002
;      CURRENT APPLICATION NUMBER: US/09/475,515A
;      CURRENT FILING DATE: 1999-12-30
;      NUMBER OF SEQ ID NOS: 90
;      SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 6
;      LENGTH: 4319
;      TYPE: DNA
;      ORGANISM: Artificial Sequence
;      FEATURE:
;      OTHER INFORMATION: Description of Artificial Sequence: synthetic
;      OTHER INFORMATION: HIV-Gag-polymerase
US-09-475-515-6

  Query Match             97.3%;  Score 58.4;  DB 4;  Length 4319;
  Best Local Similarity   98.3%;  Pred. No. 4e-10;
  Matches   59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy          1 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCGACTAGGTGGACCGGCTTCTTCAAGACC 60
              |||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||
Db        862 GACATCCGCCAGGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCTACAAGACC 921


RESULT 9
US-09-475-515-75
;  Sequence 75, Application US/09475515A
;  Patent No. 6602705
;  GENERAL INFORMATION:
;      APPLICANT: BARNETT, Susan
;      APPLICANT: ZUR MEGEDE, Jan
;      APPLICANT: SRIVASTAVA, Indresh
;      APPLICANT: LIAN, Ying
;      APPLICANT: HARTOG, Karin
;      APPLICANT: LIU, Hong
;      APPLICANT: GREER, Catherine
;      APPLICANT: SELBY, Mark
;      APPLICANT: WALKER, Christopher
;      TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;      TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;      FILE REFERENCE: 1621.002
;      CURRENT APPLICATION NUMBER: US/09/475,515A
;      CURRENT FILING DATE: 1999-12-30
;      NUMBER OF SEQ ID NOS: 90
;      SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 75
;      LENGTH: 4472
;      TYPE: DNA
;      ORGANISM: Artificial Sequence
;      FEATURE:
;      OTHER INFORMATION: Description of Artificial Sequence: synthetic
;      OTHER INFORMATION: gp160.modUS4.delV1/V2.gag.modSF2
US-09-475-515-75

  Query Match             97.3%;  Score 58.4;  DB 4;  Length 4472;
  Best Local Similarity   98.3%;  Pred. No. 4e-10;
```

Matches   59;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

```
Qy       1 GACATCCGCCAGGGCCCCCCAAGGAGGAGCCCTTCCGCGACTAGTGACCGCTTCTTCTTCAGAGACC   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3794 GACATCCGCCAGGGCCCCCCAAGGAGGAGCCCTTCCGCGACTAGTGACCGCTTCTTCTTCAGAGACC 3853
```

RESULT 10
US-09-475-515-76
; Sequence 76, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 76
; LENGTH: 4608
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modSP162.delV2.gag.modSP2
US-09-475-515-76

Query Match   97.3%;   Score 58.4;   DB 4;   Length 4608;
Best Local Similarity   98.3%;   Pred. No. 4e-10;
Matches   59;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

```
Qy       1 GACATCCGCCAGGGCCCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCTTCAAGACC   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3930 GACATCCGCCAGGGCCCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCTTCTACAAGACC 3989
```

RESULT 11
US-09-475-515-74
; Sequence 74, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 74
; LENGTH: 4689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modSP162.gag.modSP2
US-09-475-515-74

Query Match   97.3%;   Score 58.4;   DB 4;   Length 4689;
Best Local Similarity   98.3%;   Pred. No. 4e-10;
Matches   59;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

```
Qy       1 GACATCCGCCAGGGCCCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCTTCAAGACC   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4011 GACATCCGCCAGGGCCCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCTTCTACAAGACC 4070
```

RESULT 12
US-09-475-515-73
; Sequence 73, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 73
; LENGTH: 4766
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modUS4.gag.modSP2
US-09-475-515-73

Query Match   97.3%;   Score 58.4;   DB 4;   Length 4766;
Best Local Similarity   98.3%;   Pred. No. 4e-10;
Matches   59;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

```
Qy       1 GACATCCGCCAGGGCCCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCTTCAAGACC   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4088 GACATCCGCCAGGGCCCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCTTCTACAAGACC 4147
```

RESULT 13
US-09-552-950-2
; Sequence 2, Application US/09552950
; Patent No. 6541248
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; TITLE OF INVENTION: Anti-Viral Vectors
; FILE REFERENCE: 674524-2004
; CURRENT APPLICATION NUMBER: US/09/552,950
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 4307
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:gagpol-SYNgp - codon
; OTHER INFORMATION: optimised gagpol sequence
US-09-552-950-2

Query Match   79.7%;   Score 47.8;   DB 4;   Length 4307;

```
RESULT 14
US-09-552-950-5
; Sequence 5, Application US/09552950
; Patent No. 6541248
; GENERAL INFORMATION:
;   APPLICANT: Oxford Biomedica (UK) Limited
;   TITLE OF INVENTION: Anti-Viral Vectors
;   FILE REFERENCE: 674524-2004
;   CURRENT APPLICATION NUMBER: US/09/552,950
;   CURRENT FILING DATE: 2000-04-20
;   NUMBER OF SEQ ID NOS: 22
;   SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 9772
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:pSYNGP

  Qy        1 GACATCCGCGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGAC 59
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db     1957 GACATCCGCGCAAGGGCCCAAGGAACCCTTTCGCGACTACGTGGACCGGTTCTTCTACAAAAC 2015

Query Match       79.7%;  Score 47.8;  DB 4;  Length 9772;
Best Local Similarity   88.1%;  Pred. No. 1.1e-06;
Matches  52;  Conservative   0;  Mismatches   7;  Indels   0;  Gaps   0;
```

---

```
Qy        1 GACATCCGCGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGAC 59
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      850 GACATCCGCCAAGGCCCGAAGGAACCCTTTCCGGACTACGTGGACCGGTTCTACAAAAC 908

Best Local Similarity   88.1%;  Pred. No. 9.9e-07;
Matches  52;  Conservative   0;  Mismatches   7;  Indels   0;  Gaps   0;
```

---

```
RESULT 15
US-09-184-418C-100
; Sequence 100, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
;   APPLICANT: Hahn, Beatrice
;   APPLICANT: Gao, Feng
;   APPLICANT: Shaw, George
;   TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
;   TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
;   FILE REFERENCE: D6287
;   CURRENT APPLICATION NUMBER: US/09/184,418C
;   CURRENT FILING DATE: 1999-11-02
;   NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 100
;   LENGTH: 1476
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus type 1
;   FEATURE:
;   OTHER INFORMATION: isolate=94IN476.104; gene=gag
US-09-184-418C-100

Query Match       58.3%;  Score 35;  DB 4;  Length 1476;
Best Local Similarity   74.6%;  Pred. No. 0.011;
Matches  44;  Conservative   0;  Mismatches  15;  Indels   0;  Gaps   0;

Qy        1 GACATCCGCGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGAC 59
            |||||  ||||  |||||||||||| ||||| |  ||||| | ||||  ||||   ||||
Db      841 GACATAAGACAAGGGGCCCAAAAGCAACCCTTTAGAGACTAYGTAGACACCGGTTCTTTAAAAC 899
```

Search completed: May 28, 2004, 13:26:22
Job time : 6.49156 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on:          May 28, 2004, 11:33:50 ; Search time 30.7692 Seconds
                 (without alignments)
                 8870.358 Million cell updates/sec

Title:           US-09-475-704A-2
Perfect score:   60
Sequence:        1 gacatccgcaggccccaa............tggacgcgttcttcaagacc 60

Scoring table:   IDENTITY NUC
                 Gapop 10.0, Gapext 1.0

Searched:        2960401 seqs, 2274450654 residues

Total number of hits satisfying chosen parameters:     5920802

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :       Published Applications NA:*
    1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
    2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
    3:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
    4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
    5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
    6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
    7:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
    8:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
    9:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
   10:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
   11:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
   12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
   13:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
   14:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
   15:  /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
   16:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
   17:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
   18:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
   19:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | % Query Match | Length | Match DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 60 | 100.0 | 60 | 13 | US-09-899-575-2 | Sequence 2, Appli |
| 2 | 60 | 100.0 | 1509 | 10 | US-09-967-464-64 | Sequence 64, Appli |
| 3 | 60 | 100.0 | 1509 | 10 | US-09-967-464-68 | Sequence 68, Appli |
| 4 | 60 | 100.0 | 1509 | 13 | US-09-899-575-4 | Sequence 4, Appli |
| 5 | 60 | 100.0 | 1509 | 13 | US-09-899-575-21 | Sequence 21, Appli |
| 6 | 60 | 100.0 | 1608 | 15 | US-10-168-843A-27 | Sequence 27, Appl |
| 7 | 60 | 100.0 | 1914 | 15 | US-10-168-843A-29 | Sequence 29, Appl |
| 8 | 60 | 100.0 | 2493 | 15 | US-10-168-843A-31 | Sequence 31, Appl |
| 9 | 60 | 100.0 | 4350 | 15 | US-10-168-843A-33 | Sequence 33, Appl |
| 10 | 58.4 | 97.3 | 60 | 13 | US-10-387-336-20 | Sequence 20, Appl |
| 11 | 58.4 | 97.3 | 1268 | 13 | US-10-387-336-9 | Sequence 9, Appli |
| 12 | 58.4 | 97.3 | 1515 | 13 | US-10-387-336-4 | Sequence 4, Appli |
| 13 | 58.4 | 97.3 | 1853 | 13 | US-10-387-336-5 | Sequence 5, Appli |
| 14 | 58.4 | 97.3 | 1865 | 13 | US-10-387-336-78 | Sequence 78, Appl |
| 15 | 58.4 | 97.3 | 1865 | 13 | US-10-387-336-79 | Sequence 79, Appl |
| 16 | 58.4 | 97.3 | 2031 | 15 | US-10-387-336-7 | Sequence 7, Appl |
| 17 | 58.4 | 97.3 | 2799 | 15 | US-10-387-336-18 | Sequence 18, Appl |
| 18 | 58.4 | 97.3 | 2799 | 15 | US-10-241-009-53 | Sequence 53, Appl |
| 19 | 58.4 | 97.3 | 2799 | 15 | US-10-190-434B-18 | Sequence 18, Appl |
| 20 | 58.4 | 97.3 | 2799 | 15 | US-10-190-434B-53 | Sequence 53, Appl |
| 21 | 58.4 | 97.3 | 2799 | 15 | US-10-190-305A-11 | Sequence 11, Appl |
| 22 | 58.4 | 97.3 | 2799 | 15 | US-10-190-305A-74 | Sequence 74, Appl |
| 23 | 58.4 | 97.3 | 3205 | 15 | US-10-241-009-17 | Sequence 17, Appl |
| 24 | 58.4 | 97.3 | 3205 | 15 | US-10-190-434B-17 | Sequence 17, Appl |
| 25 | 58.4 | 97.3 | 3496 | 15 | US-10-241-009-15 | Sequence 15, Appl |
| 26 | 58.4 | 97.3 | 3496 | 15 | US-10-190-434B-15 | Sequence 15, Appl |
| 27 | 58.4 | 97.3 | 3564 | 15 | US-10-241-009-13 | Sequence 13, Appl |
| 28 | 58.4 | 97.3 | 3564 | 15 | US-10-241-009-14 | Sequence 14, Appl |
| 29 | 58.4 | 97.3 | 3564 | 15 | US-10-190-434B-13 | Sequence 13, Appl |
| 30 | 58.4 | 97.3 | 3564 | 15 | US-10-190-434B-14 | Sequence 14, Appl |
| 31 | 58.4 | 97.3 | 3999 | 15 | US-10-241-009-9 | Sequence 9, Appl |
| 32 | 58.4 | 97.3 | 3999 | 15 | US-10-241-009-10 | Sequence 10, Appl |
| 33 | 58.4 | 97.3 | 3999 | 15 | US-10-241-009-11 | Sequence 11, Appl |
| 34 | 58.4 | 97.3 | 3999 | 15 | US-10-190-434B-9 | Sequence 9, Appl |
| 35 | 58.4 | 97.3 | 3999 | 15 | US-10-190-434B-10 | Sequence 10, Appl |
| 36 | 58.4 | 97.3 | 3999 | 15 | US-10-190-434B-11 | Sequence 11, Appl |
| 37 | 58.4 | 97.3 | 4319 | 13 | US-10-387-336-6 | Sequence 6, Appl |
| 38 | 58.4 | 97.3 | 4472 | 13 | US-10-387-336-75 | Sequence 75, Appl |
| 39 | 58.4 | 97.3 | 4608 | 13 | US-10-387-336-76 | Sequence 76, Appl |
| 40 | 58.4 | 97.3 | 4608 | 13 | US-10-387-336-74 | Sequence 74, Appl |
| 41 | 58.4 | 97.3 | 4766 | 13 | US-10-387-336-73 | Sequence 73, Appl |
| 42 | 58.4 | 97.3 | 4773 | 16 | US-10-241-009-55 | Sequence 55, Appl |
| 43 | 58.4 | 97.3 | 4773 | 16 | US-10-241-009-16 | Sequence 16, Appl |
| 44 | 58.4 | 97.3 | 4773 | 16 | US-10-190-434B-16 | Sequence 16, Appl |
| 45 | 58.4 | 97.3 | 4773 | 15 | US-10-190-434B-55 | Sequence 55, Appl |

## ALIGNMENTS

RESULT 1
US-09-899-575-2

Sequence 2, Application US/09899575

;  Publication No. US20030223961A1
;  GENERAL INFORMATION:
;  APPLICANT: Zur Megede, Jan
;  APPLICANT: Barnett, Susan W.
;  APPLICANT: Egnelbrecht, Susan
;  APPLICANT: van Rensburg, Estrelita Janse
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP01631.102
;  CURRENT APPLICATION NUMBER: US/09/899,575
;  CURRENT FILING DATE: 2001-07-05
;  PRIOR APPLICATION NUMBER: 09/475,704
;  PRIOR FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.0
;
;  SEQ ID NO 2
;    LENGTH: 60
;    TYPE: DNA
;    ORGANISM: Human immunodeficiency virus

US-09-899-575-2

Query Match          100.0%;  Score 60;  DB 13;  Length 60;
Best Local Similarity   100.0%;  Pred. No. 2.2e-11;
Matches    60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 GACATCCGCAGGCCCCAAGGAGCCCTTCCGCGGACTAGTGACCGGCGTTCTTCAAGACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACATCCGCAGGCCCCAAGGAGCCCTTCCGCGGACTAGTGACCGGCGTTCTTCAAGACC 60

RESULT 2
US-09-967-464-64

;  Sequence 64, Application US/0967464
;  Publication No. US20030138453A1

; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Manmohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: PP16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 64
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-64

Query Match          100.0%;  Score 60;  DB 10;  Length 1509;
Best Local Similarity 100.0%;  Pred. No. 1.9e-11;
Matches  60;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 GACATCCGCCAGGGGCCCCAAGGAGACCCTTCCGCGACTAGTGAGTGGACCGGCTTCTTCAAGACC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GACATCCGCCAGGGGCCCCAAGGAGACCCTTCCGCGACTAGTGAGTGGACCGGCTTCTTCAAGACC 900

RESULT 3
US-09-967-464-68
; Sequence 68, Application US/09967464
; Publication No. US20030184531A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Manmohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: PP16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 68
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-68

Query Match          100.0%;  Score 60;  DB 10;  Length 1509;
Best Local Similarity 100.0%;  Pred. No. 1.9e-11;
Matches  60;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 GACATCCGCCAGGGGCCCCAAGGAGACCCTTCCGCGACTAGTGAGTGGACCGGCTTCTTCAAGACC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GACATCCGCCAGGGGCCCCAAGGAGACCCTTCCGCGACTAGTGAGTGGACCGGCTTCTTCAAGACC 900

RESULT 4
US-09-967-464-68

US-09-899-575-4
; Sequence 4, Application US/09899575
; Publication No. US20020223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;   POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: of HIV strain AF110967
US-09-899-575-4

Query Match          100.0%;  Score 60;  DB 13;  Length 1509;
Best Local Similarity 100.0%;  Pred. No. 1.9e-11;
Matches  60;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 GACATCCGCCAGGGGCCCCAAGGAGACCCTTCCGCGACTAGTGAGTGGACCGGCTTCTTCAAGACC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GACATCCGCCAGGGGCCCCAAGGAGACCCTTCCGCGACTAGTGAGTGGACCGGCTTCTTCAAGACC 900

RESULT 5
US-09-899-575-21
; Sequence 21, Application US/09899575
; Publication No. US20020223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;   POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: coding sequence of HIV strain AF110967
US-09-899-575-21

Query Match          100.0%;  Score 60;  DB 13;  Length 1509;
Best Local Similarity 100.0%;  Pred. No. 1.9e-11;
Matches  60;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 GACATCCGCCAGGGGCCCCAAGGAGACCCTTCCGCGACTAGTGAGTGGACCGGCTTCTTCAAGACC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GACATCCGCCAGGGGCCCCAAGGAGACCCTTCCGCGACTAGTGAGTGGACCGGCTTCTTCAAGACC 900

RESULT 6
US-10-168-843A-27

; Sequence 27, Application US/1016843A
; Publication No. US20030108562A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: International Aids Vaccine Initiative
; APPLICANT: University of Nairobi
; TITLE OF INVENTION: Improvements in or Relating to Immune Responses to HIV
; FILE REFERENCE: MJL/C1248'1/M
; CURRENT APPLICATION NUMBER: US/10/168,843A
; CURRENT FILING DATE: 2002-09-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 27
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
US-10-168-843A-27

Query Match      100.0%; Score 60; DB 15; Length 1608;
Best Local Similarity   100.0%; Pred. No. 1.9e-11;
Matches   60; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy    1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 472 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 531

RESULT 7
US-10-168-843A-29
; Sequence 29, Application US/1016843A
; Publication No. US20030108562A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: International Aids Vaccine Initiative
; APPLICANT: University of Nairobi
; TITLE OF INVENTION: Improvements in or Relating to Immune Responses to HIV
; FILE REFERENCE: MJL/C1248'1/M
; CURRENT APPLICATION NUMBER: US/10/168,843A
; CURRENT FILING DATE: 2002-09-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 29
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
US-10-168-843A-29

Query Match      100.0%; Score 60; DB 15; Length 1914;
Best Local Similarity   100.0%; Pred. No. 1.9e-11;
Matches   60; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy    1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 466 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525

Qy
Best Local Similarity   100.0%; Pred. No.; Mismatches   0;
Matches   60; Conservative   0;

Qy    1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 466 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525

RESULT 8
US-10-168-843A-31
; Sequence 31, Application US/1016843A
; Publication No. US20030108562A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: International Aids Vaccine Initiative
; APPLICANT: University of Nairobi
; TITLE OF INVENTION: Improvements in or Relating to Immune Responses to HIV
; FILE REFERENCE: MJL/C1248'1/M
; CURRENT APPLICATION NUMBER: US/10/168,843A

; CURRENT FILING DATE: 2002-09-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 31
; LENGTH: 2493
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
US-10-168-843A-31

Query Match      100.0%; Score 60; DB 15; Length 2493;
Best Local Similarity   100.0%; Pred. No. 1.9e-11;
Matches   60; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy    1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 466 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525

RESULT 9
US-10-168-843A-33
; Sequence 33, Application US/1016843A
; Publication No. US20030108562A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: International Aids Vaccine Initiative
; APPLICANT: University of Nairobi
; TITLE OF INVENTION: Improvements in or Relating to Immune Responses to HIV
; FILE REFERENCE: MJL/C1248'1/M
; CURRENT APPLICATION NUMBER: US/10/168,843A
; CURRENT FILING DATE: 2002-09-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 33
; LENGTH: 4350
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
US-10-168-843A-33

Query Match      100.0%; Score 60; DB 15; Length 4350;
Best Local Similarity   100.0%; Pred. No. 1.8e-11;
Matches   60; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy    1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 466 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525

RESULT 10
US-10-387-336-20
; Sequence 20, Application US/10387336
; Publication No. US20030223964A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/10/387,336
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/475,515A

```
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 20
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic p55
; OTHER INFORMATION: Gag Major Homology Region
US-10-387-336-20

Qy    1 GACATCCGGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACACCGCTTCTTCAAGACC  60
        |||||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACACCGCTTCTTCTACAAGACC  60

RESULT 11
US-10-387-336-9
; Sequence 9, Application US/10387336
; Publication No. US20030223964A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/10/387,336
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/475,515A
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: common region
US-10-387-336-9

Query Match             97.3%;  Score 58.4;  DB 13;  Length 1268;
Best Local Similarity   98.3%;  Pred. No. 6.8e-11;
Matches   59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    1 GACATCCGGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACACCGCTTCTTCAAGACC  60
        ||||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  862 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACACCGCTTCTTCTACAAGACC  921

RESULT 12
US-10-387-336-4
; Sequence 4, Application US/10387336
; Publication No. US20030223964A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/10/387,336
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/475,515A
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag
US-10-387-336-4

Query Match             97.3%;  Score 58.4;  DB 13;  Length 1515;
Best Local Similarity   98.3%;  Pred. No. 6.8e-11;
Matches   59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    1 GACATCCGCGCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACACCGCTTCTTCAAGACC  60
        ||||||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  862 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACACCGCTTCTTCTACAAGACC  921

RESULT 13
US-10-387-336-5
; Sequence 5, Application US/10387336
; Publication No. US20030223964A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/10/387,336
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/475,515A
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag-protease
US-10-387-336-5

Query Match             97.3%;  Score 58.4;  DB 13;  Length 1853;
Best Local Similarity   98.3%;  Pred. No. 6.7e-11;
Matches   59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    1 GACATCCGGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACACCGCTTCTTCAAGACC  60
        ||||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  862 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACACCGCTTCTTCTACAAGACC  921

RESULT 14
US-10-387-336-78
```

;
; Sequence 78, Application US/10387336
; Publication No. US20030223964A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/10/387,336
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/475,515A
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 78
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GP1
US-10-387-336-78

Query Match        97.3%;  Score 58.4;  DB 13;  Length 1865;
Best Local Similarity   98.3%;  Pred. No. 6.7e-11;
Matches  59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy     1 GACATCCGCCAGGGGCCCCAAGGAGGACCCCTTCCGCGACTAGGTGGACCGCTTCTTCAAGACC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   868 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTAGCTGGACCGCTTCTTCTACAAGACC 927

;
; Sequence 79, Application US/10387336
; Publication No. US20030223964A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/10/387,336
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/475,515A
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 79
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GP2
US-10-387-336-79

RESULT 15
US-10-387-336-79

Query Match        97.3%;  Score 58.4;  DB 13;  Length 1865;
Best Local Similarity   98.3%;  Pred. No. 6.7e-11;
Matches  59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy     1 GACATCCGCCAGGGGCCCCAAGGAGGACCCCTTCCGCGACTAGGTGGACCGCTTCTTCAAGACC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   868 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTAGCTGGACCGCTTCTTCTACAAGACC 927

Search completed: May 28, 2004, 16:09:15
Job time : 30.7692 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 28, 2004, 07:10:24 ; Search time 27.8049 Seconds
(without alignments)
9167.164 Million cell updates/sec

Title: US-09-475-704a-2
Perfect score: 60
Sequence: 1 gacatccgccagggccccaa..........tggacgcttcttcaagacc 60
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29jan04:*
1: geneseqn11980s:*
2: geneseqn11990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn20028:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn20043:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | Match DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 60 | 100.0 | 1509 | 3 | AAA51608 | Aaa51608 HIV Gag m |
| 2 | 60 | 100.0 | 1509 | 3 | AAA51610 | Aaa51610 HIV synth |
| 3 | 60 | 100.0 | 1509 | 3 | AAA51626 | Aaa51626 HIV codon |
| 4 | 60 | 100.0 | 1509 | 3 | AAL44553 | Aal44553 HIV-1 p55 |
| 5 | 60 | 100.0 | 1509 | 6 | AAL44549 | Aal44549 HIV-1 p55 |
| 6 | 60 | 100.0 | 1509 | 6 | ABL39954 | Abl39954 Synthetic |
| 7 | 60 | 100.0 | 1509 | 6 | ABL39958 | Abl39958 Synthetic |
| 8 | 60 | 100.0 | 1608 | 4 | AAD09487 | Aad09487 Human imm |
| 9 | 60 | 100.0 | 1914 | 4 | AAD09488 | Aad09488 Human imm |
| 10 | 60 | 100.0 | 2493 | 4 | AAD09489 | Aad09489 Human imm |
| 11 | 60 | 100.0 | 4350 | 4 | AAD09490 | Aad09490 Human imm |
| 12 | 58.4 | 97.3 | 60 | 3 | AAA70426 | Aaa70426 HIV p55 G |
| 13 | 58.4 | 97.3 | 1268 | 3 | AAA70417 | Aaa70417 HIV Gag c |
| 14 | 58.4 | 97.3 | 1503 | 3 | AAZ52051 | Aaz52051 Codon opt |
| 15 | 58.4 | 97.3 | 1515 | 3 | AAA70412 | Aaa70412 Synthetic |
| 16 | 58.4 | 97.3 | 1853 | 3 | AAA70413 | Aaa70413 HIV Gag-p |
| 17 | 58.4 | 97.3 | 1865 | 3 | AAA70475 | Aaa70475 HIV Gag-P |
| 18 | 58.4 | 97.3 | 1865 | 3 | AAA70476 | Aaa70476 HIV Gag-P |
| 19 | 58.4 | 97.3 | 2031 | 3 | AAA70415 | Aaa70415 Synthetic |
| 20 | 58.4 | 97.3 | 2799 | 7 | ACA03583 | Aca03583 Synthetic |
| 21 | 58.4 | 97.3 | 2799 | 7 | ACA03520 | Aca03520 Synthetic |
| 22 | 58.4 | 97.3 | 2799 | 7 | ACC78493 | Acc78493 HIV GagTa |
| 23 | 58.4 | 97.3 | 2799 | 7 | ACC78528 | Acc78528 HIV TatRe |
| 24 | 58.4 | 97.3 | 3205 | 7 | ACC78492 | Acc78492 HIV GagRT |
| 25 | 58.4 | 97.3 | 3496 | 7 | ACC78490 | Acc78490 HIV GagPr |
| 26 | 58.4 | 97.3 | 3564 | 7 | ACC78488 | Acc78488 HIV GagPo |
| 27 | 58.4 | 97.3 | 3564 | 7 | ACC78489 | Acc78489 HIV GagPo |
| 28 | 58.4 | 97.3 | 3999 | 7 | ACC78484 | Acc78484 HIV GagCo |
| 29 | 58.4 | 97.3 | 3999 | 7 | ACC78485 | Acc78485 HIV GagCo |
| 30 | 58.4 | 97.3 | 3999 | 7 | ACC78486 | Acc78486 HIV Gag-p |
| 31 | 58.4 | 97.3 | 4319 | 3 | AAA70414 | Aaa70414 HIV Gag-p |
| 32 | 58.4 | 97.3 | 4472 | 3 | AAA70472 | Aaa70472 HIV bicis |
| 33 | 58.4 | 97.3 | 4608 | 3 | AAA70473 | Aaa70473 HIV bicis |
| 34 | 58.4 | 97.3 | 4689 | 3 | AAA70471 | Aaa70471 HIV bicis |
| 35 | 58.4 | 97.3 | 4766 | 3 | AAA70470 | Aaa70470 HIV bicis |
| 36 | 58.4 | 97.3 | 4773 | 7 | ACA03519 | Aca03519 Synthetic |
| 37 | 58.4 | 97.3 | 4773 | 7 | ACA03585 | Aca03585 Synthetic |
| 38 | 58.4 | 97.3 | 4773 | 7 | ACC78491 | Acc78491 HIV GagPr |
| 39 | 58.4 | 97.3 | 4773 | 7 | ACC78530 | Acc78530 HIV TatRe |
| 40 | 58.4 | 97.3 | 5274 | 7 | ACA03518 | Aca03518 Synthetic |
| 41 | 58.4 | 97.3 | 5274 | 7 | ACC78487 | Acc78487 HIV GagCp |
| 42 | 58.4 | 97.3 | 5283 | 7 | ACA03584 | Aca03584 Synthetic |
| 43 | 58.4 | 97.3 | 5283 | 7 | ACC78529 | Acc78529 HIV TatRe |
| 44 | 58.4 | 97.3 | 6438 | 7 | ABK91605 | Abk91605 Modified |
| 45 | 58.4 | 97.3 | 8908 | 3 | AA252055 | Aa252055 Packaging |

ALIGNMENTS

RESULT 1
AAA51608
ID   AAA51608 standard; DNA; 60 BP.
XX
AC   AAA51608;
XX
DT   31-OCT-2000  (first entry)
XX
DE   HIV Gag major homology region nucleotides 841-900.
XX
KW   Gag; expression cassette; antigenic; type C; HIV; synthetic;
KW   DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS   Human immunodeficiency virus.
XX
XX
PN   WO200039304-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999; 99WO-US031273.
XX
PR   31-DEC-1998; 98US-01144495P.
PR   01-SEP-1999; 99US-0152195P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S, Zur Megede J;
XX
DR   WPI; 2000-452401/39.
XX
PT   Polynucleotide encoding antigenic type C HIV Gag polypeptide or a.HIV Env
PT   polypeptide and the polypeptide useful for immunizing a mammal especially
PT   human against HIV.
XX
PS   Claim 1; Page 92; 113pp; English.
XX
CC   Expression cassettes comprising a polynucleotide encoding antigenic type
CC   C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful
CC   in DNA immunization, generation of packaging cell lines and production of
CC   Gag- and/or Env-containing proteins. Synthetic Env and Gag expression
CC   cassettes exhibit increased potency for induction of cytotoxic T-
CC   lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-
CC   assemble into non-infectious virus-like particles which are used as a
CC   matrix for the proper presentation of an antigen entrapped or associated
CC   to the immune system of the host

SQ  Sequence 60 BP; 12 A; 24 C; 15 G; 9 T; 0 U; 0 Other;

Query Match
Best Local Similarity  100.0%;  Score 60;  DB 3;  Length 60;
Matches  60;  Conservative  0;  Pred. No. 4.5e-09;  Indels  0;  Gaps  0;

Qy  1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCGCTTCTTCAAGACC 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCGCTTCTTCAAGACC 60

RESULT 2
ID   AAA51610 standard; DNA; 1509 BP.
XX
AC   AAA51610;
XX
DT   31-OCT-2000  (first entry)
XX
DE   HIV synthetic Gag polynucleotide.
XX
KW   Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW   DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS   Human immunodeficiency virus; type C strain AF110967.
OS   Synthetic.
XX
PN   WO200039304-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;  99WO-US031273.
XX
PR   31-DEC-1998;  98US-0114495P.
PR   01-SEP-1999;  99US-0152195P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S, Zur Megede J;
XX
DR   WPI; 2000-452401/39.
XX
PT   Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env
PT   polypeptide and the polypeptide useful for immunizing a mammal especially
PT   human against HIV.
XX
PS   Claim 2; Page 93; 113pp; English.
XX
CC   Expression cassettes comprising a polynucleotide encoding antigenic type
CC   C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful
CC   in DNA immunization, generation of packaging cell lines and production of
CC   Gag- and/or Env-containing proteins. Synthetic Env and Gag expression
CC   cassettes exhibit increased potency for induction of cytotoxic T-
CC   lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-
CC   assemble into non-infectious virus-like particles which are used as a
CC   matrix for the proper presentation of an antigen entrapped or associated
CC   to the immune system of the host
XX
SQ   Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other;

Query Match
Best Local Similarity  100.0%;  Score 60;  DB 3;  Length 1509;
Matches  60;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCGCTTCTTCAAGACC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCGCTTCTTCAAGACC 900

RESULT 3
ID   AAA51626 standard; DNA; 1509 BP.
XX
AC   AAA51626;
XX
DT   31-OCT-2000  (first entry)
XX
DE   HIV codon-optimized synthetic Gag polynucleotide.
XX
KW   Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW   DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS   Human immunodeficiency virus; type C strain AF110967.
OS   Synthetic.
XX
PN   WO200039304-A2.
XX
PD   06-JUL-2000.
XX
PF   30-DEC-1999;  99WO-US031273.
XX
PR   31-DEC-1998;  98US-0114495P.
PR   01-SEP-1999;  99US-0152195P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S, Zur Megede J;
XX
DR   WPI; 2000-452401/39.
XX
PT   Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env
PT   polypeptide and the polypeptide useful for immunizing a mammal especially
PT   human against HIV.
XX
PS   Disclosure; Page 104; 113pp; English.
XX
CC   Expression cassettes comprising a polynucleotide encoding antigenic type
CC   C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful
CC   in DNA immunization, generation of packaging cell lines and production of
CC   Gag- and/or Env-containing proteins. Synthetic Env and Gag expression
CC   cassettes exhibit increased potency for induction of cytotoxic T-
CC   lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-
CC   assemble into non-infectious virus-like particles which are used as a
CC   matrix for the proper presentation of an antigen entrapped or associated
CC   to the immune system of the host
XX
SQ   Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 U; 0 Other;

Query Match
Best Local Similarity  100.0%;  Score 60;  DB 3;  Length 1509;
Matches  60;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCGCTTCTTCAAGACC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCGCTTCTTCAAGACC 900

RESULT 4
ID   AAL44553 standard; DNA; 1509 BP.
XX
AC   AAL44553;
XX
DT   29-AUG-2003  (revised)
DT   08-NOV-2002  (first entry)
XX
DE   HIV-1 p55gag polypeptide coding sequence 4.
XX
KW   HIV; dsi; vaccine; gene; immune response; microparticle;
KW   adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
KW   polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
KW   submicron emulsion; viral infection; bacterial infection;
KW   parasitic infection; HIV-1 p55gag polypeptide.
XX
OS   Human immunodeficiency virus 1.
XX

| FH | Key | Location/Qualifiers |
| --- | --- | --- |
| FT | variation | replace(282, C) |
| FT | variation | /*tag= a |
| FT | variation | replace(331, T) |
| FT | variation | /*tag= b |
| FT | variation | replace(332, C) |
| FT | variation | /*tag= c |
| FT | variation | replace(549, G) |
| FT | variation | /*tag= d |
| FT | variation | replace(553, T) |
| FT | variation | /*tag= e |
| FT | variation | replace(783, G) |
| FT | variation | /*tag= f |
| FT | variation | replace(816, G) |
| FT | variation | /*tag= g |
| FT | variation | replace(999, T) |
| FT | variation | /*tag= h |
| FT | variation | replace(1002, C) |
| FT | variation | /*tag= i |
| FT | variation | replace(1089, G) |
| FT | variation | /*tag= j |
| FT | variation | replace(1149, G) |
| FT | variation | /*tag= k |
| FT | variation | replace(1158, C) |
| FT | variation | /*tag= l |
| XX | | |
| PN | WO20226209-A2. |
| XX | |
| PD | 04-APR-2002. |
| XX | |
| PF | 28-SEP-2001; 2001WO-US030540. |
| XX | |
| PR | 28-SEP-2000; 2000US-0236105P. |
| PR | 30-AUG-2001; 2001US-0315905P. |
| XX | |
| PA | (CHIR ) CHIRON CORP. |
| XX | |
| PI | O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M; |
| PI | Ulmer J, Dubensky TW; |
| XX | |
| DR | WPI: 2002-519084/55. |
| XX | |
| PS | Claim 72; Fig 6; 100pp; English. |
| XX | |

Claim 72; Fig 6; 100pp; English.

The invention relates to a method of raising an immune response in a host animal. The method of the invention comprises administering a microparticle that has an adsorbent surface to which a first biologically active macromolecule (e.g. a nucleic acid) has been adsorbed. The microparticle comprises a polymer microparticle of poly(alpha-hydroxy acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester, a polycyanoacrylate, a detergent, and submicron emulsion. The method is useful for immunising a host animal against viral, bacterial or parasitic infections. The present DNA sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to standardise OS field)

| SQ | Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 U; 0 Other; |

| Qy | 1 | GACATCCGCCAGGGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGGCTTCTTCAAGACC | 60 |
| Db | 841 | GACATCCGCCAGGGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGGCTTCTTCAAGACC | 900 |

---

Query Match
Best Local Similarity 100.0%; Score 60; DB 6; Length 1509;
Matches 60; Conservative 0; Pred. No. 5.7e-09;
　　　　　　　　　　Mismatches 0; Indels 0; Gaps 0;

---

| ID | AAL44549 standard; DNA; 1509 BP. |
| XX | |
| AC | AAL44549; |
| XX | |
| DT | 29-AUG-2003 (revised) |
| DT | 08-NOV-2002 (first entry) |
| XX | |
| DE | HIV-1 p55gag polypeptide coding sequence 2. |
| XX | |
| KW | HIV; ds; vaccine; gene; immune response; microparticle; |
| KW | adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid; |
| KW | polycaprolactone; polyorthoester; polycyanoacrylate; detergent; |
| KW | submicron emulsion; viral infection; bacterial infection; |
| KW | parasitic infection; HIV-1 p55gag polypeptide. |
| XX | |
| OS | Human immunodeficiency virus 1. |
| XX | |
| PN | WO20226209-A2. |
| XX | |
| PD | 04-APR-2002. |
| XX | |
| PF | 28-SEP-2001; 2001WO-US030540. |
| XX | |
| PR | 28-SEP-2000; 2000US-0236105P. |
| PR | 30-AUG-2001; 2001US-0315905P. |
| XX | |
| PA | (CHIR ) CHIRON CORP. |
| XX | |
| PI | O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M; |
| PI | Ulmer J, Dubensky TW; |
| XX | |
| DR | WPI: 2002-519084/55. |
| XX | |
| PS | Claim 72; Fig 2; 100pp; English. |

Claim 72; Fig 2; 100pp; English.

The invention relates to a method of raising an immune response in a host animal. The method of the invention comprises administering a microparticle that has an adsorbent surface to which a first biologically active macromolecule (e.g. a nucleic acid) has been adsorbed. The microparticle comprises a polymer microparticle of poly(alpha-hydroxy acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester, a polycyanoacrylate, a detergent, and submicron emulsion. The method is useful for immunising a host animal against viral, bacterial or parasitic infections. The present DNA sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to standardise OS field)

| SQ | Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other; |

| Qy | 1 | GACATCCGCCAGGGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGGCTTCTTCAAGACC | 60 |
| Db | 841 | GACATCCGCCAGGGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGACCGGCTTCTTCAAGACC | 900 |

---

| ID | ABL39954 standard; DNA; 1509 BP. |
| XX | |
| AC | ABL39954; |
| XX | |
| DT | 15-MAY-2002 (first entry) |
| XX | |
| DE | Synthetic Gag polynucleotide sequence SEQ ID NO:4. |

Query Match
Best Local Similarity 100.0%; Score 60; DB 6; Length 1509;
Matches 60; Conservative 0; Pred. No. 5.7e-09;
　　　　　　　　　　Mismatches 0; Indels 0; Gaps 0;

AAL44549

XX
KW  Human immunodeficiency virus type C; antigenic HIV type C protein;
KW  immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW  immunostimulant; gene therapy; gene; ds.
XX
PN  Human immunodeficiency virus; type C.
OS  Synthetic.
XX
PR  05-JUL-2000; 2000US-00610313.
XX
PA  (CHIR ) CHIRON CORP.
PA  (UYST-) UNIV STELLENBOSCH.
XX
DR  WO20020494493-A2.
XX
PD  17-JAN-2002.
XX
PP  05-JUL-2001; 2001WO-US021241.
XX
PI  Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg BJ;
XX
WPI; 2002-154920/20.
XX
DR  New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT  applications including DNA immunization or generation of packaging cell
PT  lines, particularly in gene therapy.
XX
PS  Example 1; Fig 2; 233pp; English.
XX
CC  The present invention describes expression cassettes comprising a
CC  polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC  type C polypeptides. The expression cassettes comprise any of the HIV
CC  type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC  (I). (I) have immunostimulant activity and can be used in gene therapy.
CC  The HIV type C polynucleotides are useful in applications including DNA
CC  immunisation, generation of packaging cell lines, and production of HIV
CC  Type C proteins. The polynucleotides are particularly useful in gene
CC  therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC  ABB06204 to ABB06215 represent sequences used in the exemplification of
CC  the present invention
XX
SQ  Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 U; 0 Other;

Query Match         100.0%;  Score 60;  DB 6;  Length 1509;
Best Local Similarity   100.0%;  Pred. No. 5.7e-09;
Matches  60;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy           1 GACATCCGCCAGGGGCCCCAAGGAGGACCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 GACATCCGCCAGGGGCCCCAAGGAGGACCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 900

RESULT 7
ID  ABL39958
AC  ABL39958 standard; DNA; 1509 BP.
XX
DT  15-MAY-2002  (first entry)
XX
DB  Synthetic Gag polynucleotide sequence SEQ ID NO:21.
XX
Db  Human immunodeficiency virus type C; antigenic HIV type C protein;
KW  immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW  immunostimulant; gene therapy; gene; ds.
XX
OS  Human immunodeficiency virus; type C.
OS  Synthetic.
XX
PN  WO20020494493-A2.
XX
PD  17-JAN-2002.

---

PP  05-JUL-2001; 2001WO-US021241.
XX
PR  05-JUL-2000; 2000US-00610313.
XX
PA  (CHIR ) CHIRON CORP.
PA  (UYST-) UNIV STELLENBOSCH.
XX
PI  Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg BJ;
XX
WPI; 2002-154920/20.
XX
PT  New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT  applications including DNA immunization or generation of packaging cell
PT  lines, particularly in gene therapy.
XX
PS  Example 1; Fig 6; 233pp; English.
XX
CC  The present invention describes expression cassettes comprising a
CC  polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC  type C polypeptides. The expression cassettes comprise any of the HIV
CC  type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC  (I). (I) have immunostimulant activity and can be used in gene therapy.
CC  The HIV type C polynucleotides are useful in applications including DNA
CC  immunisation, generation of packaging cell lines, and production of HIV
CC  Type C proteins. The polynucleotides are particularly useful in gene
CC  therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC  ABB06204 to ABB06215 represent sequences used in the exemplification of
CC  the present invention
XX
SQ  Sequence 1509 BP; 321 A; 560 C; 470 G; 158 T; 0 U; 0 Other;

Query Match         100.0%;  Score 60;  DB 6;  Length 1509;
Best Local Similarity   100.0%;  Pred. No. 5.7e-09;
Matches  60;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy           1 GACATCCGCCAGGGGCCCCAAGGAGGACCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 GACATCCGCCAGGGGCCCCAAGGAGGACCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 900

RESULT 8
ID  AAD09487
AC  AAD09487 standard; DNA; 1608 BP.
XX
DT  10-SEP-2001  (first entry)
XX
DB  Human immunodeficiency virus A (HIV A) gene.
XX
KW  Human immunodeficiency virus A; HIV A; immunogen; anti-HIV; vaccine;
KW  gene therapy; fusion protein; modified vaccinia virus Ankara vector; MVA;
KW  cytotoxic T-lymphocyte; CTL; epitope; ds.
XX
OS  Human immunodeficiency virus.
XX
FH  Key              Location/Qualifiers
FT  CDS              19..1602
FT                   /tag= a
FT                   /product= "HIV A immunogen"
XX
PN  WO200147955-A2.
XX
PD  05-JUL-2001.
XX
PP  22-DEC-2000; 2000WO-GB004984.
XX
PR  23-DEC-1999;  99GB-00030294.
PR  14-OCT-2000; 2000GB-00025234.
XX
PA  (MEDI-) MEDICAL RES COUNCIL.
PA  (ITAI-) INT AIDS VACCINE INITIATIVE.
PA  (UYNA-) UNIV NAIROBI.

XX
PI   Hanke T, Mcmichael AJ;
XX
DR   WPI; 2001-418221/44.
DR   P-PSDB; AAB04825.
XX
PT   Novel immunogen for stimulating anti-HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T-lymphocyte epitopes of HIV
PT   protein.
XX
PS   Claim 29; Fig 2A; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
CC   prevent N-terminal myristoylation; and a synthetic polypeptide comprising
CC   human cytotoxic T-lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV-specific immune response in
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject. The
CC   invention also relates to method of stimulating anti-HIV immune response
CC   in a human subject which comprises administering one or more times an
CC   amount of nucleic acid molecule sufficient to prime an immune response to
CC   the immunogen, or else may be packaged within a delivery means, such as a
CC   modified vaccinia virus Ankara (MVA) to boost the immune response to
CC   common portion of the immunogens. The present DNA sequence encodes human
CC   immunodeficiency virus A immunogen (HIV A) fusion protein construct. HIV
CC   A immunogen consists of about 73% of gag protein fused to a string of 25
CC   partially overlapping human CTL epitopes. The gag domain of HIV A
CC   contains p24 and p17 in an order reversed to the viral gag p17-p24-p15
CC   polyprotein.
XX
SQ   Sequence 1608 BP; 314 A; 598 C; 458 G; 238 T; 0 U; 0 Other;

Qy          1 GACATCCGCCAGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GACATCCGCCAGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 531

Query Match            100.0%;  Score 60;  DB 4;  Length 1608;
Best Local Similarity  100.0%;  Pred. No. 5.7e-09;
Matches      60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

RESULT 9
ID   AAD09488 standard; DNA; 1914 BP.
XX
AC   AAD09488;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human immunodeficiency virus TA (HIV TA) gene.
XX
KW   Human immunodeficiency virus TA; HIV TA; immunogen; anti-HIV; vaccine;
KW   gene therapy; fusion protein; modified vaccinia virus Ankara vector; MVA;
KW   cytotoxic T-lymphocyte; CTL; epitope; ds.
XX
OS   Human immunodeficiency virus.
XX
FH   Key               Location/Qualifiers
FT   CDS               13..1914
FT                     /*tag= a
FT                     /product= "HIV TA immunogen"
XX
PN   WO200147955-A2.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-GB004984.
XX
PR   23-DEC-1999; 99GB-00030294.
PR   14-OCT-2000; 2000GB-00025234.

---

XX
PA   (MEDI-) MEDICAL RES COUNCIL.
PA   (ITAI-) INT AIDS VACCINE INITIATIVE.
PA   (UYNA-) UNIV NAIROBI.
XX
PI   Hanke T, Mcmichael AJ;
XX
DR   WPI; 2001-418221/44.
DR   P-PSDB; AAE04826.
XX
PT   Novel immunogen for stimulating anti-HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T-lymphocyte epitopes of HIV
PT   protein.
XX
PS   Example 2; Fig 6B; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from HIV clade, parts of p17 and p24, modified to
CC   prevent N-terminal myristoylation; and a synthetic polypeptide comprising
CC   human cytotoxic T-lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV-specific immune response in
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject. The
CC   invention also relates to method of stimulating anti-HIV immune response
CC   in a human subject which comprises administering one or more times an
CC   amount of nucleic acid molecule sufficient to prime an immune response to
CC   the immunogen, or else may be packaged within a delivery means, such as a
CC   modified vaccinia virus Ankara (MVA) to boost the immune response to
CC   common portion of the immunogens. The present DNA sequence encodes human
CC   immunodeficiency virus TA immunogen (HIV TA) fusion protein construct.
CC   HIV TA construct is based on HIV A immunogen. The HIV A immunogen
CC   consists of about 73% of gag protein fused to a string of 25 partially
CC   overlapping human CTL epitopes. The gag domain of HIV A contains p24 and
CC   p17 in an order reversed to the viral gag p17-p24-p15 polyprotein. HIV TA
CC   shares the same design rationale with HIV A construct, but additionally
CC   includes the HIV-1 clade A tat sequence positioned between the gag
CC   protein and synthetic human CTL epitopes
XX
SQ   Sequence 1914 BP; 383 A; 728 C; 534 G; 269 T; 0 U; 0 Other;

Qy          1 GACATCCGCCAGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 GACATCCGCCAGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525

Query Match            100.0%;  Score 60;  DB 4;  Length 1914;
Best Local Similarity  100.0%;  Pred. No. 5.8e-09;
Matches      60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

RESULT 10
ID   AAD09489 standard; DNA; 2493 BP.
XX
AC   AAD09489;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human immunodeficiency virus AeT (HIV AeT) gene.
XX
KW   Human immunodeficiency virus AeT; HIV AeT; immunogen; anti-HIV; vaccine;
KW   gene therapy; fusion protein; modified vaccinia virus Ankara vector; MVA;
KW   cytotoxic T-lymphocyte; CTL; epitope; ds.
XX
OS   Human immunodeficiency virus.
XX
FH   Key               Location/Qualifiers
FT   CDS               2179..2493
FT                     /*tag= a
FT                     /product= "HIV AeT immunogen"
XX
PN   WO200147955-A2.

XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-GB004984.
XX
PR   23-DEC-1999;  99GB-0030294.
PR   14-OCT-2000; 2000GB-0025234.
XX
PN
PA   (MEDI-) MEDICAL RES COUNCIL.
PA   (ITAI-) INT AIDS VACCINE INITIATIVE.
PA   (UYNA-) UNIV NAIROBI.
XX
PI   Hanke T, Mcmichael AJ;
XX
DR   WPI; 2001-418221/44.
DR   P-PSDB; AAB04827.
XX
PT   Novel immunogen for stimulating anti-HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T-lymphocyte epitopes of HIV
PT   protein.
XX
PS   Example 2; Fig 7B; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
CC   prevent N-terminal myristoylation; and a synthetic polypeptide comprising
CC   human cytotoxic T-lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV-specific immune response in
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject. The
CC   invention also relates to method of stimulating anti-HIV immune response
CC   in a human subject which comprises administering one or more times an
CC   amount of nucleic acid molecule sufficient to prime an immune response to
CC   the immunogen, or else may be packaged within a delivery means, such as a
CC   modified vaccinia virus Ankara (MVA) to boost the immune response to
CC   common portion of the immunogens. The present DNA sequence encodes human
CC   immunodeficiency virus AeT immunogen (HIV AeT) fusion protein construct.
CC   HIV AeT construct is based on HIV A immunogen. The HIV A immunogen
CC   consists of about 73% of gag protein fused to a string of 25 partially
CC   overlapping human CTL epitopes. The gag domain of HIV A contains p24 and
CC   p17 in an order reversed to the viral gag p17-p24-p15 polyprotein. HIV
CC   AeT shares the same design rationale with HIV A construct, but
CC   additionally includes the HIV-1 clade A tat sequence which is present on
CC   the same construct, but expressed as a separate polypeptide, by virtue of
CC   the inclusion of an internal ribosome entry site
XX
SQ   Sequence 2493 BP; 516 A; 887 C; 685 G; 405 T; 0 U; 0 Other;

Query Match             100.0%;  Score 60;  DB 4;  Length 2493;
Best Local Similarity   100.0%;  Pred. No. 5.9e-09;
Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

  Qy          1 GACATCCGGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
                ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db        466 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525

RESULT 11
ID   AAD09490 standard; DNA; 4350 BP.
AC   AAD09490;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human immunodeficiency virus PPA immunogen.
XX
KW   human immunodeficiency virus; PPA; immunogen; anti-HIV; vaccine;
KW   gene therapy; fusion protein; modified vaccinia virus Ankara vector; MVA;
KW   cytotoxic T-lymphocyte; CTL; epitope; ds.
XX

OS   Human immunodeficiency virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             13. 4350
FT                   /*tag= a
FT                   /product= "PPA immunogen"
XX
PN   WO20147955-A2.
XX
PD   05-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-GB004984.
XX
PR   23-DEC-1999;  99GB-0030294.
PR   14-OCT-2000; 2000GB-0025234.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
PA   (ITAI-) INT AIDS VACCINE INITIATIVE.
PA   (UYNA-) UNIV NAIROBI.
XX
PI   Hanke T, Mcmichael AJ;
XX
DR   WPI; 2001-418221/44.
DR   P-PSDB; AAB04828.
XX
PT   Novel immunogen for stimulating anti-HIV immune response, has a portion
PT   of gag protein of HIV from HIV clade, parts of p17, p24 and synthetic
PT   polypeptide comprising human cytotoxic T-lymphocyte epitopes of HIV
PT   protein.
XX
PS   Claim 40; Fig 8B; 65pp; English.
XX
CC   The invention relates to human immunodeficiency virus immunogens and
CC   their corresponding DNA molecules. An immunogen comprises a portion of
CC   gag protein of HIV from an HIV clade, parts of p17 and p24, modified to
CC   prevent N-terminal myristoylation; and a synthetic polypeptide comprising
CC   human cytotoxic T-lymphocyte (CTL) epitopes of HIV protein. This
CC   immunogen is designed to elicit an HIV-specific immune response in
CC   humans. The immunogen is useful in the preparation of a medicament such
CC   as vaccine to prevent or treat HIV infection in a human subject. The
CC   invention also relates to method of stimulating anti-HIV immune response
CC   in a human subject which comprises administering one or more times an
CC   amount of nucleic acid molecule sufficient to prime an immune response to
CC   the immunogen, or else may be packaged within a delivery means, such as a
CC   modified vaccinia virus Ankara (MVA) to boost the immune response to
CC   the immunogens. The present DNA sequence encodes human
CC   immunodeficiency virus PPA immunogen fusion protein construct. PPA
CC   construct is based on HIV A immunogen. The HIV A immunogen consists of
CC   about 73% of gag protein fused to a string of 25 partially overlapping
CC   human CTL epitopes. The gag domain of HIV A contains p24 and p17 in an
CC   order reversed to the viral gag p17-p24-p15 polyprotein. PPA shares the
CC   same design rationale with HIV A construct, but additionally includes the
CC   HIV-1 clade A tat sequence, C-pol protein, N-pol protein, rev and nef
CC   proteins
XX
SQ   Sequence 4350 BP; 888 A; 1630 C; 1259 G; 573 T; 0 U; 0 Other;

Query Match             100.0%;  Score 60;  DB 4;  Length 4350;
Best Local Similarity   100.0%;  Pred. No. 6.1e-09;
Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

  Qy          1 GACATCCGGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
                ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
  Db        466 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 525

RESULT 12
ID   AAA70426 standard; DNA; 60 BP.
AC   AAA70426;
XX
DT   15-SEP-2003  (revised)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 28, 2004, 09:05:41 ; Search time 246.567 Seconds
(without alignments)
10547.179 Million cell updates/sec

Title: US-09-475-704A-2
Perfect score: 60
Sequence: 1 gaactccgccaggccccaa...........tggaccgcttcttcaagacc 60

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 2167151b995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: gb_ba:*
16: gb_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 60 | 100.0 | 60 | 6 | AX455886 | AX455886 Sequence |
| 2 | 60 | 100.0 | 1509 | 6 | AX455888 | AX455888 Sequence |
| 3 | 60 | 100.0 | 1509 | 6 | AX455905 | AX455905 Sequence |
| 4 | 60 | 100.0 | 1509 | 6 | AX468544 | AX468544 Sequence |
| 5 | 60 | 100.0 | 1509 | 6 | AX468548 | AX468548 Sequence |
| 6 | 60 | 100.0 | 1608 | 6 | AX188558 | AX188558 Sequence |
| 7 | 60 | 100.0 | 1914 | 6 | AX188560 | AX188560 Sequence |
| 8 | 60 | 100.0 | 2493 | 6 | AX188562 | AX188562 Sequence |
| 9 | 60 | 100.0 | 4350 | 6 | AX188564 | AX188564 Sequence |
| 10 | 58.4 | 97.3 | 60 | 6 | BD263651 | BD263651 Improved |
| 11 | 58.4 | 97.3 | 60 | 6 | AR373334 | AR373334 Sequence |
| 12 | 58.4 | 97.3 | 1268 | 6 | BD263642 | BD263642 Improved |
| 13 | 58.4 | 97.3 | 1268 | 6 | AR373325 | AR373325 Sequence |
| 14 | 58.4 | 97.3 | 1509 | 12 | AF201927 | AF201927 Synthetic |
| 15 | 58.4 | 97.3 | 1515 | 6 | BD263637 | BD263637 Improved |
| 16 | 58.4 | 97.3 | 1515 | 6 | AR373320 | AR373320 Sequence |
| 17 | 58.4 | 97.3 | 1847 | 6 | AR373325 | AR373325 Sequence |
| 18 | 58.4 | 97.3 | 1853 | 12 | AF202464 | AF202464 Synthetic |
| 19 | 58.4 | 97.3 | 1853 | 6 | BD263638 | BD263638 Improved |
| 20 | 58.4 | 97.3 | 1853 | 6 | AR373321 | AR373321 Sequence |
| 21 | 58.4 | 97.3 | 1865 | 6 | BD263700 | BD263700 Improved |
| 22 | 58.4 | 97.3 | 1865 | 6 | BD263701 | BD263701 Improved |
| 23 | 58.4 | 97.3 | 1865 | 6 | AR373383 | AR373383 Sequence |
| 24 | 58.4 | 97.3 | 1865 | 6 | AR373384 | AR373384 Sequence |
| 25 | 58.4 | 97.3 | 2031 | 6 | BD263640 | BD263640 Improved |
| 26 | 58.4 | 97.3 | 2031 | 6 | AR373323 | AR373323 Sequence |
| 27 | 58.4 | 97.3 | 4319 | 6 | BD263639 | BD263639 Improved |
| 28 | 58.4 | 97.3 | 4319 | 6 | AR373322 | AR373322 Sequence |
| 29 | 58.4 | 97.3 | 4472 | 6 | BD263697 | BD263697 Improved |
| 30 | 58.4 | 97.3 | 4472 | 6 | AR373380 | AR373380 Sequence |
| 31 | 58.4 | 97.3 | 4608 | 6 | BD263698 | BD263698 Improved |
| 32 | 58.4 | 97.3 | 4608 | 6 | AR373381 | AR373381 Sequence |
| 33 | 58.4 | 97.3 | 4689 | 6 | BD263696 | BD263696 Improved |
| 34 | 58.4 | 97.3 | 4689 | 6 | AR373379 | AR373379 Sequence |
| 35 | 58.4 | 97.3 | 4766 | 6 | BD263695 | BD263695 Improved |
| 36 | 58.4 | 97.3 | 4766 | 6 | AR373378 | AR373378 Sequence |
| 37 | 58.4 | 97.3 | 6438 | 6 | AX427919 | AX427919 Sequence |
| 38 | 58.4 | 97.3 | 9167 | 6 | AX427933 | AX427933 Sequence |
| 39 | 58.4 | 97.3 | 9170 | 6 | AX427928 | AX427928 Sequence |
| 40 | 58.4 | 97.3 | 9189 | 6 | AX427921 | AX427921 Sequence |
| 41 | 58.4 | 97.3 | 9194 | 6 | AX427926 | AX427926 Sequence |
| 42 | 58.4 | 97.3 | 9194 | 6 | AX427925 | AX427925 Sequence |
| 43 | 58.4 | 97.3 | 9407 | 6 | AX427937 | AX427937 Sequence |
| 44 | 58.4 | 97.3 | 9782 | 6 | AX427935 | AX427935 Sequence |
| 45 | 58.4 | 97.3 | 9783 | 6 | AX427934 | AX427934 Sequence |

ALIGNMENTS

RESULT 1
AX455886
LOCUS        AX455886        60 bp    DNA    linear    PAT    06-JUL-2002
DEFINITION   Sequence 2 from Patent WO0204493.
ACCESSION    AX455886
VERSION      AX455886.1    GI:21714879
KEYWORDS
SOURCE       Human immunodeficiency virus
  ORGANISM   Human immunodeficiency virus
             Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
             lentivirus group.
REFERENCE    1
  AUTHORS    zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE      Polynucleotides encoding antigenic hiv type c polypeptides,
             polypeptides and uses thereof

Pred. No. is the number of results predicted by chance to have a

JOURNAL          Patent: WO 0204493-A 2 17-JAN-2002;
                 CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
     source          1..60
                     /organism="Human immunodeficiency virus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:12721"
ORIGIN

Qy          Query Match          100.0%;     Score 60;   DB 6;   Length 60;
            Best Local Similarity   100.0%;   Pred. No. 9.7e-07;
            Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

          1 GACATCCGCCAGGGCCCCAAGGAGCCCCTTCCGCGGACTACGTGGGACCGGCTTCTTCTTCAAGACC   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          1 GACATCCGCCAGGGCCCCAAGGAGCCCCTTCCGCGGACTACGTGGGACCGGCTTCTTCTTCAAGACC   60

RESULT 2
LOCUS          AX455888          1509 bp          DNA          linear          PAT 06-JUL-2002
DEFINITION     Sequence 4 from Patent WO0204493.
ACCESSION      AX455888
VERSION        AX455888.1  GI:21714881
KEYWORDS
SOURCE         synthetic construct
  ORGANISM     synthetic construct
               artificial sequences.
REFERENCE      1
  AUTHORS      zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE        Polynucleotides encoding antigenic hiv type c polypeptides,
               polypeptides and uses thereof
  JOURNAL      Patent: WO 0204493-A 4 17-JAN-2002;
               CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
     source          1..1509
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="synthetic Gag of HIV strain AF110967"
ORIGIN

Qy          Query Match          100.0%;     Score 60;   DB 6;   Length 1509;
            Best Local Similarity   100.0%;   Pred. No. 5.6e-07;
            Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

          1 GACATCCGCCAGGGCCCCAAGGAGCCCCTTCCGCGGACTACGTGGGACCGGCTTCTTCTTCAAGACC   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        841 GACATCCGCCAGGGCCCCAAGGAGCCCCTTCCGCGGACTACGTGGGACCGGCTTCTTCTTCAAGACC  900

RESULT 3
LOCUS          AX455905          1509 bp          DNA          linear          PAT 06-JUL-2002
DEFINITION     Sequence 21 from Patent WO0204493.
ACCESSION      AX455905
VERSION        AX455905.1  GI:21714897
KEYWORDS
SOURCE         synthetic construct
  ORGANISM     synthetic construct
               artificial sequences.
REFERENCE      1
  AUTHORS      zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE        Polynucleotides encoding antigenic hiv type c polypeptides,
               polypeptides and uses thereof
  JOURNAL      Patent: WO 0204493-A 21 17-JAN-2002;
               CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
     source          1..1509
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="synthetic Gag coding sequence of HIV strain

---

ORIGIN          AF110967"

Qy          Query Match          100.0%;     Score 60;   DB 6;   Length 1509;
            Best Local Similarity   100.0%;   Pred. No. 5.6e-07;
            Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

          1 GACATCCGCCAGGGCCCCAAGGAGCCCCTTCCGCGGACTACGTGGGACCGGCTTCTTCTTCAAGACC   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        841 GACATCCGCCAGGGCCCCAAGGAGCCCCTTCCGCGGACTACGTGGGACCGGCTTCTTCTTCAAGACC  900

RESULT 4
LOCUS          AX468544          1509 bp          DNA          linear          PAT 16-JUL-2002
DEFINITION     Sequence 64 from Patent WO0226209.
ACCESSION      AX468544
VERSION        AX468544.1  GI:21901374
KEYWORDS
SOURCE         Human immunodeficiency virus 1 (HIV-1)
  ORGANISM     Human immunodeficiency virus 1
               Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
               lentivirus group.
REFERENCE      1
  AUTHORS      O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
               Singh,M., Ulmer,J. and Dubensky,T.W.
  TITLE        Microparticles for delivery of the heterologous nucleic acids
  JOURNAL      Patent: WO 0226209-A 64 04-APR-2002;
               CHIRON CORPORATION (US)
FEATURES
     source          1..1509
                     /organism="Human immunodeficiency virus 1"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11676"
ORIGIN

Qy          Query Match          100.0%;     Score 60;   DB 6;   Length 1509;
            Best Local Similarity   100.0%;   Pred. No. 5.6e-07;
            Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

          1 GACATCCGCCAGGGCCCCAAGGAGCCCCTTCCGCGGACTACGTGGGACCGGCTTCTTCTTCAAGACC   60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        841 GACATCCGCCAGGGCCCCAAGGAGCCCCTTCCGCGGACTACGTGGGACCGGCTTCTTCTTCAAGACC  900

RESULT 5
LOCUS          AX468548          1509 bp          DNA          linear          PAT 16-JUL-2002
DEFINITION     Sequence 68 from Patent WO0226209.
ACCESSION      AX468548
VERSION        AX468548.1  GI:21901378
KEYWORDS
SOURCE         Human immunodeficiency virus 1 (HIV-1)
  ORGANISM     Human immunodeficiency virus 1
               Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
               lentivirus group.
REFERENCE      1
  AUTHORS      O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
               Singh,M., Ulmer,J. and Dubensky,T.W.
  TITLE        Microparticles for delivery of the heterologous nucleic acids
  JOURNAL      Patent: WO 0226209-A 68 04-APR-2002;
               CHIRON CORPORATION (US)
FEATURES
     source          1..1509
                     /organism="Human immunodeficiency virus 1"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11676"
ORIGIN  .

Qy          Query Match          100.0%;     Score 60;   DB 6;   Length 1509;
            Best Local Similarity   100.0%;   Pred. No. 5.6e-07;
            Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
RESULT 6
LOCUS       AX188558               1608 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 27 from Patent WO0147955.
ACCESSION   AX188558
VERSION     AX188558.1  GI:15142199
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Hanke,T.M. and Mcmichael,A.J.
  TITLE     Improvements in or relating to immune responses to hiv
  JOURNAL   Patent: WO 0147955-A 27 05-JUL-2001;
            MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine
            Initiative (US) ; University of Nairobi (KE)
FEATURES             Location/Qualifiers
     source          1..1608
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Chimeric polynucleotide"
ORIGIN

Qy          1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 900

Query Match  100.0%;  Score 60;  DB 6;  Length 1608;
Best Local Similarity  100.0%;  Pred. No. 5.5e-07;
Matches 60;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
RESULT 7
LOCUS       AX188560               1914 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 29 from Patent WO0147955.
ACCESSION   AX188560
VERSION     AX188560.1  GI:15142200
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Hanke,T.M. and Mcmichael,A.J.
  TITLE     Improvements in or relating to immune responses to hiv
  JOURNAL   Patent: WO 0147955-A 29 05-JUL-2001;
            MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine
            Initiative (US) ; University of Nairobi (KE)
FEATURES             Location/Qualifiers
     source          1..1914
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Chimeric polynucleotide"
ORIGIN

Qy          1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 531

Query Match  100.0%;  Score 60;  DB 6;  Length 1914;
Best Local Similarity  100.0%;  Pred. No. 5.4e-07;
Matches 60;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
RESULT 8
LOCUS       AX188562               2493 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 31 from Patent WO0147955.
ACCESSION   AX188562
VERSION     AX188562.1  GI:15142201
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Hanke,T.M. and Mcmichael,A.J.
  TITLE     Improvements in or relating to immune responses to hiv
  JOURNAL   Patent: WO 0147955-A 31 05-JUL-2001;
            MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine
            Initiative (US) ; University of Nairobi (KE)
FEATURES             Location/Qualifiers
     source          1..2493
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Chimeric polynucleotide"
ORIGIN

Qy          1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 525

Query Match  100.0%;  Score 60;  DB 6;  Length 2493;
Best Local Similarity  100.0%;  Pred. No. 5.1e-07;
Matches 60;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
RESULT 9
LOCUS       AX188564               4350 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 33 from Patent WO0147955.
ACCESSION   AX188564
VERSION     AX188564.1  GI:15142202
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Hanke,T.M. and Mcmichael,A.J.
  TITLE     Improvements in or relating to immune responses to hiv
  JOURNAL   Patent: WO 0147955-A 33 05-JUL-2001;
            MEDICAL RESEARCH COUNCIL (GB) ; International Aids Vaccine
            Initiative (US) ; University of Nairobi (KE)
FEATURES             Location/Qualifiers
     source          1..4350
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Chimeric polynucleotide"
ORIGIN

Qy          1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 525

Query Match  100.0%;  Score 60;  DB 6;  Length 4350;
Best Local Similarity  100.0%;  Pred. No. 4.7e-07;
Matches 60;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
RESULT 10
LOCUS       BD263651                 60 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Improved expression of HIV polypeptides and production of
            virus-like particles.
ACCESSION   BD263651
VERSION     BD263651.1  GI:33073419
KEYWORDS    JP 2002533124-A/18.
SOURCE      synthetic construct
```

ORGANISM   synthetic construct
           artificial sequences.
REFERENCE  1 (bases 1 to 60)
AUTHORS    Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
           Greer,C., Selby,M. and Walker,C.
TITLE      Improved expression of HIV polypeptides and production of
           virus-like particles
JOURNAL    Patent: JP 2002533124-A 18 08-OCT-2002;
           CHIRON CORP
COMMENT    OS   Artificial Sequence
           PN   JP 2002533124-A/18
           PD   08-OCT-2002
           PF   30-DEC-1999 JP 2000591193
           PR   31-DEC-1998 US 60/114495,01-DEC-1999 US   60/168471 PI
           SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI
           KARIN HARTOG,
           PI   HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC
           C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
           C12N5/10,
           PC   C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC
           Description of Artificial Sequence: synthetic p55 Gag CC Major
           Homology Region
           FH   Key             Location/Qualifiers
           FT   source          1..60
                                /organism="synthetic construct"
                                /mol_type="genomic DNA"
                                /db_xref="taxon:32630"

ORIGIN
Query Match             97.3%;  Score 58.4;  DB 6;  Length 60;
Best Local Similarity   98.3%;  Pred. No. 2.7e-06;
Matches   59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy          1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60

RESULT 11
LOCUS       AR373334      60 bp    DNA      linear   PAT 18-DEC-2003
DEFINITION  Sequence 20 from patent US 6602705.
ACCESSION   AR373334
VERSION     AR373334.1  GI:40075437
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 60)
AUTHORS     Barnett,S.W., Megede,J., Greer,C. and Selby,M.
TITLE       Expression of HIV polypeptides and production of virus-like
            particles
JOURNAL     Patent: US 6602705-A 20 05-AUG-2003;
FEATURES             Location/Qualifiers
     source          1..60
                     /organism="unknown"
                     /mol_type="genomic DNA"

ORIGIN
Query Match             97.3%;  Score 58.4;  DB 6;  Length 60;
Best Local Similarity   98.3%;  Pred. No. 2.7e-06;
Matches   59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy          1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60

RESULT 12
BD263642

---

LOCUS       BD263642    1268 bp    DNA      linear   PAT 17-JUL-2003
DEFINITION  Improved expression of HIV polypeptides and production of
            virus-like particles.
ACCESSION   BD263642
VERSION     BD263642.1  GI:33073410
KEYWORDS    JP 2002533124-A/9.
SOURCE      synthetic construct
ORGANISM    synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1268)
AUTHORS     Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
            Greer,C., Selby,M. and Walker,C.
TITLE       Improved expression of HIV polypeptides and production of
            virus-like particles
JOURNAL     Patent: JP 2002533124-A 9 08-OCT-2002;
            CHIRON CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2002533124-A/9
            PD   08-OCT-2002
            PF   30-DEC-1999 JP 2000591193
            PR   31-DEC-1998 US 60/114495,01-DEC-1999 US   60/168471 PI
            SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI
            KARIN HARTOG,
            PI   HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC
            C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
            C12N5/10,
            PC   C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC
            Description of Artificial Sequence: synthetic Gag common CC
            region
            FH   Key             Location/Qualifiers
            FT   source          1..1268
                                 /organism="synthetic construct"
                                 /mol_type="genomic DNA"
                                 /db_xref="taxon:32630"

ORIGIN
Query Match             97.3%;  Score 58.4;  DB 6;  Length 1268;
Best Local Similarity   98.3%;  Pred. No. 1.6e-06;
Matches   59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy          1 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        862 GACATCCGCCAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC 921

RESULT 13
AR373325
LOCUS       AR373325    1268 bp    DNA      linear   PAT 18-DEC-2003
DEFINITION  Sequence 9 from patent US 6602705.
ACCESSION   AR373325
VERSION     AR373325.1  GI:4007S428
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1268)
AUTHORS     Barnett,S.W., Megede,J., Greer,C. and Selby,M.
TITLE       Expression of HIV polypeptides and production of virus-like
            particles
JOURNAL     Patent: US 6602705-A 9 05-AUG-2003;
FEATURES             Location/Qualifiers
     source          1..1268
                     /organism="unknown"
                     /mol_type="genomic DNA"

ORIGIN
Query Match             97.3%;  Score 58.4;  DB 6;  Length 1268;
Best Local Similarity   98.3%;  Pred. No. 1.6e-06;
Matches   59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

```
Qy      1 GACATCCGGCCAGGGGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACACGGCCTTCTTCAAGACC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 GACATCCGGCCAGGGGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACACGGCCTTCTTCAAGACC 921
```

RESULT 14

LOCUS           AF201927                1509 bp    DNA     linear   SYN 16-MAR-2000
DEFINITION      Synthetic construct gag protein gene, complete cds.
ACCESSION       AF201927
VERSION         AF201927.1  GI:7248702
KEYWORDS        .
SOURCE          synthetic construct
  ORGANISM      synthetic construct
                artificial sequences.
REFERENCE       1  (bases 1 to 1509)
  AUTHORS       zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,
                Selby,M., Otten,G.R. and Barnett,S.W.
  TITLE         Increased expression and immunogenicity of sequence-modified human
                immunodeficiency virus type 1 gag gene
  JOURNAL       J. Virol. 74 (6), 2628-2635 (2000)
   MEDLINE      20148954
    PUBMED      10684277
REFERENCE       2  (bases 1 to 1509)
  AUTHORS       zur Megede,J. and Barnett,S.W.
  TITLE         Direct Submission
  JOURNAL       Submitted (04-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton,
                Emeryville, CA 94608, USA
FEATURES             Location/Qualifiers
     source          1..1509
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     CDS             1..1509
                     /codon_start=1
                     /transl_table=11
                     /product="gag protein"
                     /protein_id="AAF43628.1"
                     /db_xref="GI:7248703"
                     /note="HIV-1SF2 p55gag"
                     /translation="MGARASVLSGGELDKWEKIRLRPGGKKKYKLKHIVWASRELERF
                     AVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIDVKDTKEALE
                     KIEEEQNKSKKKAQQAAAAAGTNSSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKV
                     VEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVH
                     PVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIV
                     RMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKT
                     ILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNPATIMMQRGNFRNQRKTV
                     KCFNCGKEGHTARNCRAPRKKGCWRCGREGHQMKDCTERQANFLGKIWPSYKGRPGNF
                     LQSRPEPTAPPEESFRFGEEKTTPSQKQEPIDKELYPLTSLRSLFGNDPSSQ"

ORIGIN

Qy      1 GACATCCGGCCAGGGGCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACACGGCCTTCTTCAAGACC 60
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    856 GACATCCGGCCAGGGGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACACGGCCTTCTTCAAGACC 915

RESULT 15

LOCUS           BD263637                1515 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION      Improved expression of HIV polypeptides and production of
                virus-like particles.
ACCESSION       BD263637
VERSION         BD263637.1  GI:33073405
KEYWORDS        JP 2002533124-A/4.
SOURCE          synthetic construct
  ORGANISM      synthetic construct
                artificial sequences.
REFERENCE       1  (bases 1 to 1515)
  AUTHORS       Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
```

---

TITLE           Improved expression of HIV polypeptides and production of
                virus-like particles
JOURNAL         Patent: JP 2002533124-A 4 08-OCT-2002;
                CHIRON CORP
COMMENT         OS   Artificial Sequence
                PN   JP 200253124-A/4
                PD   08-OCT-2002
                PF   30-DEC-1999 JP 2000591193
                PR   31-DEC-1998 US   60/114495,01-DEC-1999 US   60/168471 PI
                SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI
                KARIN HARTOG,
                PI   HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC
                C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
                C12N5/10,
                PC   C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC
                Description of Artificial Sequence: synthetic HIV-Gag FH
                Description of Artificial Sequence: synthetic HIV-Gag Key
FEATURES             Location/Qualifiers
     source          1..1515
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     FT   source          1..1515
     FT                   /organism='Artificial Sequence'

ORIGIN

Qy      1 GACATCCGGCCAGGGGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACACGGCCTTCTTCAAGACC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 GACATCCGGCCAGGGGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACACGGCCTTCTTCAAGACC 921

```
Query Match           97.3%;  Score 58.4;  DB 6;  Length 1515;
Best Local Similarity 98.3%;  Pred. No. 1.6e-06;
Matches  59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;
```

Search completed: May 28, 2004, 13:23:16
Job time : 247.567 secs

---

TITLE           Greer,C., Selby,M. and Walker,C.
                Improved expression of HIV polypeptides and production of
                virus-like particles
JOURNAL         Patent: JP 2002533124-A 4 08-OCT-2002;

```
Query Match           97.3%;  Score 58.4;  DB 12;  Length 1509;
Best Local Similarity 98.3%;  Pred. No. 1.6e-06;
Matches  59;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;
```

OM nucleic - nucleic search, using sw model

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

Run on:          May 28, 2004, 11:33:50 ; Search time 30.7692 Seconds
                                          (without alignments)
                                          8870.358 Million cell updates/sec

Title:           US-09-475-704A-1
Perfect score:   60
Sequence:        1 gacatcaagcagggcccaa..........tggaccgcttcttcaagacc 60

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2960401 seqs, 2274450654 residues

Total number of hits satisfying chosen parameters:    5920802

Post-processing: Minimum Match 0%
Maximum DB seq length: 200000000
Minimum DB seq length: 0
                 Listing first 45 summaries

Database :       Published Applications NA:*
                 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                 6: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
                 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
                 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
                 11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
                 13: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
                 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
                 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
                 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                 17: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
                 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 60 | 100.0 | 60 | 13 | US-09-899-575-1 | Sequence 1, Appl |
| 2 | 60 | 100.0 | 60 | 13 | US-09-899-575-53 | Sequence 53, Appl |
| 3 | 60 | 100.0 | 1479 | 10 | US-09-967-464-63 | Sequence 63, Appl |
| 4 | 60 | 100.0 | 1479 | 10 | US-09-967-464-67 | Sequence 67, Appl |
| 5 | 60 | 100.0 | 1479 | 13 | US-09-899-575-3 | Sequence 3, Appli |
| 6 | 60 | 100.0 | 1479 | 13 | US-09-899-575-20 | Sequence 20, Appl |
| 7 | 60 | 100.0 | 1491 | 13 | US-09-899-575-99 | Sequence 99, Appl |
| 8 | 60 | 100.0 | 1494 | 13 | US-09-899-575-51 | Sequence 51, Appl |
| 9 | 60 | 100.0 | 2742 | 15 | US-10-190-435-20 | Sequence 20, Appl |
| 10 | 60 | 100.0 | 2742 | 15 | US-10-190-435-57 | Sequence 57, Appl |
| 11 | 60 | 100.0 | 2742 | 15 | US-11-190-305A-15 | Sequence 15, Appl |
| 12 | 60 | 100.0 | 2742 | 15 | US-10-190-305A-81 | Sequence 81, Appl |
| 13 | 60 | 100.0 | 3162 | 15 | US-10-190-435-18 | Sequence 18, Appl |
| 14 | 60 | 100.0 | 3462 | 15 | US-10-190-435-16 | Sequence 16, Appl |
| 15 | 60 | 100.0 | 60 | 13 | US-10-190-435-13 | Sequence 13, Appl |
| 16 | 60 | 100.0 | 3537 | 15 | US-10-190-435-14 | Sequence 14, Appl |
| 17 | 60 | 100.0 | 3537 | 15 | US-10-190-435-15 | Sequence 15, Appl |
| 18 | 60 | 100.0 | 3930 | 15 | US-10-190-435-9 | Sequence 9, Appli |
| 19 | 60 | 100.0 | 3930 | 15 | US-10-190-435-10 | Sequence 10, Appl |
| 20 | 60 | 100.0 | 3930 | 15 | US-10-190-435-11 | Sequence 11, Appl |
| 21 | 60 | 100.0 | 4419 | 15 | US-10-190-435-19 | Sequence 19, Appl |
| 22 | 60 | 100.0 | 4419 | 15 | US-10-190-305A-14 | Sequence 14, Appl |
| 23 | 60 | 100.0 | 4483 | 15 | US-10-190-435-35 | Sequence 35, Appl |
| 24 | 60 | 100.0 | 4606 | 15 | US-10-190-435-34 | Sequence 34, Appl |
| 25 | 60 | 100.0 | 4615 | 15 | US-10-190-435-36 | Sequence 36, Appl |
| 26 | 60 | 100.0 | 4702 | 15 | US-10-190-435-38 | Sequence 38, Appl |
| 27 | 60 | 100.0 | 4713 | 15 | US-10-190-435-59 | Sequence 59, Appl |
| 28 | 60 | 100.0 | 4713 | 15 | US-10-190-305A-83 | Sequence 83, Appl |
| 29 | 60 | 100.0 | 4716 | 15 | US-10-190-435-17 | Sequence 17, Appl |
| 30 | 60 | 100.0 | 4716 | 15 | US-10-190-435-12 | Sequence 12, Appl |
| 31 | 60 | 100.0 | 5145 | 15 | US-10-190-305A-13 | Sequence 13, Appl |
| 32 | 60 | 100.0 | 5145 | 15 | US-10-190-435-12 | Sequence 12, Appl |
| 33 | 60 | 100.0 | 5145 | 15 | US-10-190-305A-12 | Sequence 12, Appl |
| 34 | 60 | 100.0 | 5184 | 15 | US-10-190-435-58 | Sequence 58, Appl |
| 35 | 60 | 100.0 | 9166 | 13 | US-10-359-120-82 | Sequence 82, Appl |
| 36 | 60 | 100.0 | 9169 | 13 | US-10-359-120-168 | Sequence 168, App |
| 37 | 60 | 100.0 | 9785 | 13 | US-10-359-120-169 | Sequence 169, App |
| 38 | 60 | 100.0 | 9785 | 13 | US-10-359-120-176 | Sequence 176, App |
| 39 | 58.4 | 97.3 | 9788 | 13 | US-10-359-120-174 | Sequence 174, App |
| 40 | 58.4 | 97.3 | 1092 | 15 | US-10-138-098-12 | Sequence 12, Appl |
| 41 | 58.4 | 97.3 | 1092 | 15 | US-10-138-098-13 | Sequence 13, Appl |
| 42 | 58.4 | 97.3 | 3020 | 15 | US-10-138-098-21 | Sequence 21, Appl |
| 43 | 58.4 | 97.3 | 3021 | 15 | US-10-138-098-22 | Sequence 22, Appl |
| 44 | 58.4 | 97.3 | 3021 | 15 | US-10-138-098-23 | Sequence 23, Appl |
| 45 | 55.2 | 92.0 | 3021 | 15 | US-10-138-098-24 | Sequence 24, Appl |
|  | 60 | 100.0 | 60 | 13 | US-09-899-575-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1
US-09-899-575-1

; Sequence 1, Application US/0989575
; Publication No. US20030239961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede,
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 1
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus
US-09-899-575-1

  Query Match       100.0%;  Score 60;  DB 13;  Length 60;
  Best Local Similarity   100.0%;
  Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1 GACATCAAGCAGGGCCCCAAGGAGACCCTTCCGGACTACGTGGACCGCTTCTTCAAGACC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GACATCAAGCAGGGCCCCAAGGAGACCCTTCCGGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 2
US-09-899-575-53
; Sequence 53, Application US/0989575
; Publication No. US20030239961A1

; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 53
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV Type C Gag
; OTHER INFORMATION: Major Homology Region Optimized
US-09-899-575-53

RESULT 3
US-09-967-464-63
; Sequence 63, Application US/09967464
; Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Manohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: PP16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 63
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-63

Query Match           100.0%;  Score 60;  DB 13;  Length 60;
Best Local Similarity 100.0%;  Pred. No. 2.5e-12;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1 GACATCAAGCAGGCAGGGCCCCCAAGGAGCCCTTCCGCGACTAGTGGACCGCTTCTTCAAGACC 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GACATCAAGCAGGCAGGGCCCCCAAGGAGCCCTTCCGCGACTAGTGGACCGCTTCTTCAAGACC 60

Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Manohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: PP16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 67
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-67

RESULT 5
US-09-899-575-3
; Sequence 3, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
US-09-899-575-3

Query Match           100.0%;  Score 60;  DB 10;  Length 1479;
Best Local Similarity 100.0%;  Pred. No. 2.6e-12;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1 GACATCAAGCAGGCAGGGCCCCCAAGGAGCCCTTCCGCGACTAGTGGACCGCTTCTTCAAGACC 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       844 GACATCAAGCAGGCAGGGCCCCCAAGGAGCCCTTCCGCGACTAGTGGACCGCTTCTTCAAGACC 903

RESULT 4
US-09-967-464-67
; Sequence 67, Application US/09967464

Qy         1 GACATCAAGCAGGCAGGGCCCCCAAGGAGCCCTTCCGCGACTAGTGGACCGCTTCTTCAAGACC 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       844 GACATCAAGCAGGCAGGGCCCCCAAGGAGCCCTTCCGCGACTAGTGGACCGCTTCTTCAAGACC 903

RESULT 6
US-09-899-575-20

Query Match           100.0%;  Score 60;  DB 13;  Length 1479;
Best Local Similarity 100.0%;  Pred. No. 2.6e-12;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

; OTHER INFORMATION: Description of HIV strain AF110965
US-09-899-575-3

; Sequence 20, Application US/0989995575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
;   APPLICANT: Zur Megede, Jan
;   APPLICANT: Barnett, Susan W.
;   APPLICANT: Egnelbrecht, Susan
;   APPLICANT: van Rensburg, Estrelita Janse
;   TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP01631.102
;   CURRENT APPLICATION NUMBER: US/09/899,575
;   CURRENT FILING DATE: 2001-07-05
;   PRIOR APPLICATION NUMBER: 09/475,704
;   PRIOR FILING DATE: 1999-12-30
;   NUMBER OF SEQ ID NOS: 135
;   SOFTWARE: Patentin Ver. 2.0
;   SEQ ID NO 20
;   LENGTH: 1479
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
US-09-899-575-20

Query Match       100.0%;  Score 60;  DB 13;  Length 1479;
Best Local Similarity  100.0%;  Pred. No. 2.6e-12;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 GACATCAAGCAGCCCCAAGGAGAGCCCTTCCGCGACTAGGTGGACCGCTTCTTCTTCAAGACC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      844 GACATCAAGCAGCCCCAAGGAGAGCCCTTCCGCGACTAGGTGGACCGCTTCTTCTTCAAGACC 903

RESULT 7
US-09-899-575-99
; Sequence 99, Application US/0989995575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
;   APPLICANT: Zur Megede, Jan
;   APPLICANT: Barnett, Susan W.
;   APPLICANT: Egnelbrecht, Susan
;   APPLICANT: van Rensburg, Estrelita Janse
;   TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP01631.102
;   CURRENT APPLICATION NUMBER: US/09/899,575
;   CURRENT FILING DATE: 2001-07-05
;   PRIOR APPLICATION NUMBER: 09/475,704
;   PRIOR FILING DATE: 1999-12-30
;   NUMBER OF SEQ ID NOS: 135
;   SOFTWARE: Patentin Ver. 2.0
;   SEQ ID NO 99
;   LENGTH: 1491
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Gag_TV2_C_ZAopt
US-09-899-575-99

Query Match       100.0%;  Score 60;  DB 13;  Length 1491;
Best Local Similarity  100.0%;  Pred. No. 2.6e-12;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 GACATCAAGCAGGGGCCCCAAGGAGGAGCCCTTCGGCGGCGACTAGGTGGACCGCTTCTTCTTCAAGACC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GACATCAAGCAGGGGCCCCAAGGAGGAGCCCTTCGGCGGCGACTAGGTGGACCGCTTCTTCTTCAAGACC 900

RESULT 8
US-09-899-575-51
; Sequence 51, Application US/0989995575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
;   APPLICANT: Zur Megede, Jan
;   APPLICANT: Barnett, Susan W.
;   APPLICANT: Egnelbrecht, Susan
;   APPLICANT: van Rensburg, Estrelita Janse
;   TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP01631.102
;   CURRENT APPLICATION NUMBER: US/09/899,575
;   CURRENT FILING DATE: 2001-07-05
;   PRIOR APPLICATION NUMBER: 09/475,704
;   PRIOR FILING DATE: 1999-12-30
;   NUMBER OF SEQ ID NOS: 135
;   SOFTWARE: Patentin Ver. 2.0
;   SEQ ID NO 51
;   LENGTH: 1494
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: HIV Type C
;   OTHER INFORMATION: Gag optimized
US-09-899-575-51

Query Match       100.0%;  Score 60;  DB 13;  Length 1494;
Best Local Similarity  100.0%;  Pred. No. 2.6e-12;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 GACATCAAGCAGGGGCCCCAAGGAGGAGCCCTTCGGCGACTAGGTGGACCGCTTCTTCTTCAAGACC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GACATCAAGCAGGGGCCCCAAGGAGGAGCCCTTCGGCGACTAGGTGGACCGCTTCTTCTTCAAGACC 900

RESULT 9
US-10-190-435-20
; Sequence 20, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan W.
;   APPLICANT: LIAN, Ying
;   APPLICANT: ENGELBRECHT, Susan
;   APPLICANT: VAN RENSBURG, Estrelita J.
;   TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP18133.003 / 2302-18133
;   CURRENT APPLICATION NUMBER: US/10/190,435
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 319
;   SOFTWARE: Patentin Ver. 2.0
;   SEQ ID NO 20
;   LENGTH: 2742
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
US-10-190-435-20

Query Match       100.0%;  Score 60;  DB 15;  Length 2742;
Best Local Similarity  100.0%;  Pred. No. 2.7e-12;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCGGCGACTAGGTGGACCGCTTCTTCTTCAAGACC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      850 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCGGCGACTAGGTGGACCGCTTCTTCTTCAAGACC 909

RESULT 10
US-10-190-435-57
; Sequence 57, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan W.

RESULT 11
US-10-190-305A-15
Sequence 15, Application US/10190305A
Publication No. US20030198621A1
GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 15
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
US-10-190-305A-15

Qy    1 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCCGCGGACTACGTGGACCGGCTTCTTCAAGACC 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2104 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCCGCGGACTACGTGGACCGGCTTCTTCAAGACC 2163

Query Match         100.0%;  Score 60;  DB 15;  Length 2742;
Best Local Similarity   100.0%;  Pred. No. 2.7e-12;
Matches  60;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

RESULT 12
US-10-305A-81
Sequence 81, Application US/1019030SA
Publication No. US20030198621A1
GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93

---

RESULT 13
US-10-190-435-18
Sequence 18, Application US/10190435
Publication No. US20030143248A1
GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 18
; LENGTH: 3162
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmut_C
US-10-190-435-18

Qy    1 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCCGCGGACTACGTGGACCGGCTTCTTCAAGACC 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  850 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCCGCGGACTACGTGGACCGGCTTCTTCAAGACC 909

Query Match         100.0%;  Score 60;  DB 15;  Length 3162;
Best Local Similarity   100.0%;  Pred. No. 2.7e-12;
Matches  60;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

RESULT 14
US-10-190-435-16
Sequence 16, Application US/10190435
Publication No. US20030143248A1
GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 16
; LENGTH: 3462
; TYPE: DNA
; ORGANISM: Artificial Sequence

```
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: GagProtinaRTmut_C
US-10-190-435-16

    Query Match         100.0%;  Score 60;  DB 15;  Length 3462;
    Best Local Similarity  100.0%;  Pred. No. 2.7e-12;
    Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 GACATCAAGCAGGAGGGGCCCCAAGGAGCAGCCCTTCCGCGACTAGCGTGGACCCGCTTCTTCAAGACC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    850 GACATCAAGCAGGAGGGGCCCCAAGGAGCAGCCCTTCCGCGACTAGCGTGGACCCGCTTCTTCAAGACC 909


RESULT 15
US-10-190-435-13
Sequence 13, Application US/10190435
Publication No. US20030143248A1
;   GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan W.
;   APPLICANT: LIAN, Ying
;   APPLICANT: ENGELBRECHT, Susan
;   APPLICANT: VAN RENSBURG, Estrelita J.
;   TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP18133.003 / 2302-18133
;   CURRENT APPLICATION NUMBER: US/10/190,435
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 319
;   SOFTWARE: Patentin Ver. 2.0
;   SEQ ID NO 13
;   LENGTH: 3531
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: GagPolmut_C
US-10-190-435-13

    Query Match         100.0%;  Score 60;  DB 15;  Length 3531;
    Best Local Similarity  100.0%;  Pred. No. 2.7e-12;
    Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 GACATCAAGCAGGAGGGGCCCCAAGGAGCAGCCCTTCCGCGACTAGCGTGGACCCGCTTCTTCAAGACC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    850 GACATCAAGCAGGAGGGGCCCCAAGGAGCAGCCCTTCCGCGACTAGCGTGGACCCGCTTCTTCAAGACC 909
```

Search completed: May 28, 2004, 16:09:15
Job time : 31.7692 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 28, 2004, 09:05:41 ; Search time 246.567 Seconds
(without alignments)
10547.179 Million cell updates/sec

Title:        US-09-475-704A-1
Perfect score: 60
Sequence:     1 gacatcaaggcagggccccaa..........tggaccgcttcttcaagacc 60
Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0
Searched:     3470272 seqs, 2167151699S residues
Total number of hits satisfying chosen parameters:        6940544
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  GenEmbl:*

```
1:  gb_ba:*
2:  gb_htg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_htg:*
17: em_in:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_sy:*
29: em_un:*
30: em_vi:*
31: em_htg_hum:*
32: em_htg_inv:*
33: em_htg_other:*
34: em_htg_mus:*
35: em_htg_pln:*
36: em_htg_rod:*
37: em_htg_mam:*
38: em_htg_vrt:*
39: em_sy:*
40: em_htgo_hum:*
41: em_htgo_mus:*
42: em_htgo_other:*
```

## SUMMARIES

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 60 | 100.0 | 60 | 6 | AX455885 | AX455885 Sequence |
| 2 | 60 | 100.0 | 60 | 6 | AX455937 | AX455937 Sequence |
| 3 | 60 | 100.0 | 1479 | 6 | AX455887 | AX455887 Sequence |
| 4 | 60 | 100.0 | 1479 | 6 | AX455904 | AX455904 Sequence |
| 5 | 60 | 100.0 | 1479 | 6 | AX468543 | AX468543 Sequence |
| 6 | 60 | 100.0 | 1479 | 6 | AX468547 | AX468547 Sequence |
| 7 | 60 | 100.0 | 1485 | 12 | AY181195 | AY181195 Synthetic |
| 8 | 60 | 100.0 | 1491 | 6 | AX455983 | AX455983 Sequence |
| 9 | 60 | 100.0 | 1494 | 6 | AX455935 | AX455935 Sequence |
| 10 | 60 | 100.0 | 1494 | 6 | AX427930 | AX427930 Sequence |
| 11 | 60 | 100.0 | 9166 | 6 | AX427931 | AX427931 Sequence |
| 12 | 60 | 100.0 | 9169 | 6 | AX427938 | AX427938 Sequence |
| 13 | 60 | 100.0 | 9785 | 6 | AX427936 | AX427936 Sequence |
| 14 | 60 | 100.0 | 9788 | 6 | AX427936 | AX427936 Sequence |
| 15 | 55.2 | 92.0 | 1509 | 6 | AX455886 | AX455886 Sequence |
| 16 | 55.2 | 92.0 | 1509 | 6 | AX455888 | AX455888 Sequence |
| 17 | 55.2 | 92.0 | 1509 | 6 | AX455905 | AX455905 Sequence |
| 18 | 55.2 | 92.0 | 1608 | 6 | AX468544 | AX468544 Sequence |
| 19 | 55.2 | 92.0 | 1608 | 6 | AX468548 | AX468548 Sequence |
| 20 | 55.2 | 92.0 | 1914 | 6 | AX188558 | AX188558 Sequence |
| 21 | 55.2 | 92.0 | 2493 | 12 | AX188560 | AX188560 Sequence |
| 22 | 55.2 | 92.0 | 4350 | 6 | AX188562 | AX188562 Sequence |
| 23 | 55.2 | 89.3 | 60 | 6 | AX188564 | AX188564 Sequence |
| 24 | 53.6 | 89.3 | 60 | 6 | BD263651 | BD263651 Improved |
| 25 | 53.6 | 89.3 | 1268 | 6 | BD263642 | BD263642 Improved |
| 26 | 53.6 | 89.3 | 1268 | 6 | BD263642 | BD263642 Improved |
| 27 | 53.6 | 89.3 | 1509 | 12 | AP201927 | AP201927 Synthetic |
| 28 | 53.6 | 89.3 | 1515 | 6 | BD263637 | BD263637 Improved |
| 29 | 53.6 | 89.3 | 1515 | 6 | AR373320 | AR373320 Sequence |
| 30 | 53.6 | 89.3 | 1847 | 12 | AR373320 | AR373320 Sequence |
| 31 | 53.6 | 89.3 | 1847 | 12 | AF202464 | AF202464 Synthetic |
| 32 | 53.6 | 89.3 | 1853 | 6 | AF202465 | AF202465 Synthetic |
| 33 | 53.6 | 89.3 | 1853 | 6 | AR373321 | AR373321 Improved |
| 34 | 53.6 | 89.3 | 1865 | 6 | BD263700 | BD263700 Improved |
| 35 | 53.6 | 89.3 | 1865 | 6 | BD263701 | BD263701 Improved |
| 36 | 53.6 | 89.3 | 1865 | 6 | AR373383 | AR373383 Sequence |
| 37 | 53.6 | 89.3 | 1865 | 6 | AR373384 | AR373384 Sequence |
| 38 | 53.6 | 89.3 | 2031 | 6 | BD263640 | BD263640 Improved |
| 39 | 53.6 | 89.3 | 2031 | 6 | BD263638 | BD263638 Improved |
| 40 | 53.6 | 89.3 | 4319 | 6 | BD263639 | BD263639 Improved |
| 41 | 53.6 | 89.3 | 4319 | 6 | AR373322 | AR373322 Improved |
| 42 | 53.6 | 89.3 | 4472 | 6 | AR373323 | AR373323 Improved |
| 43 | 53.6 | 89.3 | 4472 | 6 | BD263697 | BD263697 Improved |
| 44 | 53.6 | 89.3 | 4608 | 6 | BD263698 | BD263698 Improved |
| 45 | 53.6 | 89.3 | 4608 | 6 | AR373381 | AR373381 Sequence |

## ALIGNMENTS

RESULT 1
AX455885
```
LOCUS       AX455885              60 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 1 from Patent WO0204493.
ACCESSION   AX455885
VERSION     AX455885.1  GI:21714878
KEYWORDS    .
SOURCE      Human immunodeficiency virus
  ORGANISM  Human immunodeficiency virus
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
```

Pred. No. is the number of results predicted by chance to have a

JOURNAL     Patent: WO 0204493-A 1 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
    source  Location/Qualifiers
            1..60
            /organism="Human immunodeficiency virus"
            /mol_type="unassigned DNA"
            /db_xref="taxon:12721"
ORIGIN

Qy    Query Match       100.0%;   Score 60; DB 6; Length 60;
      Best Local Similarity  100.0%;   Pred. No. 2.8e-07;
      Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Db  1 GACATCAAGCAGGGCCCCAAGGAGGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
    1 GACATCAAGCAGGGCCCCAAGGAGGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 2
LOCUS       AX455937                60 bp    DNA      linear    PAT 06-JUL-2002
DEFINITION  Sequence 53 from Patent WO0204493.
ACCESSION   AX455937
VERSION     AX455937.1   GI:21714921
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 53 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
    source  Location/Qualifiers
            1..60
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="HIV Type C Gag Major Homology Region Optimized"
ORIGIN

Qy    Query Match       100.0%;   Score 60; DB 6; Length 60;
      Best Local Similarity  100.0%;   Pred. No. 2.8e-07;
      Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Db  1 GACATCAAGCAGGGCCCCAAGGAGGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
    1 GACATCAAGCAGGGCCCCAAGGAGGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 3
LOCUS       AX455887                1479 bp   DNA      linear    PAT 06-JUL-2002
DEFINITION  Sequence 3 from Patent WO0204493.
ACCESSION   AX455887
VERSION     AX455887.1   GI:21714880
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 3 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
    source  Location/Qualifiers
            1..1479
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="synthetic Gag of HIV strain AF110965"

ORIGIN

Qy    Query Match       100.0%;   Score 60; DB 6; Length 1479;
      Best Local Similarity  100.0%;   Pred. No. 1.9e-07;
      Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Db 844 GACATCAAGCAGGGCCCCAAGGAGGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 903

RESULT 4
LOCUS       AX455904                1479 bp   DNA      linear    PAT 06-JUL-2002
DEFINITION  Sequence 20 from Patent WO0204493.
ACCESSION   AX455904
VERSION     AX455904.1   GI:21714896
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 20 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
    source  Location/Qualifiers
            1..1479
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="synthetic Gag coding sequence of HIV strain
            AF110965"
ORIGIN

Qy    Query Match       100.0%;   Score 60; DB 6; Length 1479;
      Best Local Similarity  100.0%;   Pred. No. 1.9e-07;
      Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Db 844 GACATCAAGCAGGGCCCCAAGGAGGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 903

RESULT 5
LOCUS       AX468543                1479 bp   DNA      linear    PAT 16-JUL-2002
DEFINITION  Sequence 63 from Patent WO0226209.
ACCESSION   AX468543
VERSION     AX468543.1   GI:21901373
KEYWORDS    .
SOURCE      Human immunodeficiency virus 1 (HIV-1)
  ORGANISM  Human immunodeficiency virus 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1
  AUTHORS   O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S...
            Singh,M., Ulmer,J. and Dubensky,T.W.
  TITLE     Microparticles for delivery of the heterologous nucleic acids
  JOURNAL   Patent: WO 0226209-A 63 04-APR-2002;
            CHIRON CORPORATION (US)
FEATURES
    source  Location/Qualifiers
            1..1479
            /organism="Human immunodeficiency virus 1"
            /mol_type="unassigned DNA"
            /db_xref="taxon:11676"
ORIGIN

Qy    Query Match       100.0%;   Score 60; DB 6; Length 1479;
      Best Local Similarity  100.0%;   Pred. No. 1.9e-07;
      Matches   60;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 6

Qy    1 GACATCAAGCAGGAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
Db  844 GACATCAAGCAGGAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 903

LOCUS        AX468547              1479 bp    DNA     linear   PAT 16-JUL-2002
DEFINITION   Sequence 67 from Patent WO0226209.
ACCESSION    AX468547
VERSION      AX468547.1  GI:21901377
KEYWORDS     .
SOURCE       Human immunodeficiency virus 1 (HIV-1)
  ORGANISM   Human immunodeficiency virus 1
             Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
             lentivirus group.
REFERENCE    1
  AUTHORS    O'Hagan,D., Otten,G., Donnelly,J.J., Polo,J.M., Barnett,S.,
             Singh,M., Ulmer,J. and Dubensky,T.W.
  TITLE      Microparticles for delivery of the heterologous nucleic acids
  JOURNAL    Patent: WO 0226209-A 67 04-APR-2002;
             CHIRON CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1479
                     /organism="Human immunodeficiency virus 1"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11676"
ORIGIN

RESULT 7

Qy    1 GACATCAAGCAGGAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
Db  844 GACATCAAGCAGGAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 903

LOCUS        AY181195              1485 bp    DNA     linear   SYN 02-OCT-2003
DEFINITION   Synthetic construct HIV-1-derived gag protein (gag) gene, complete
             cds.
ACCESSION    AY181195
VERSION      AY181195.1  GI:37413992
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1  (bases 1 to 1485)
  AUTHORS    Gao,F., Li,Y., Decker,J.M., Peyerl,F.W., Bibollet-Ruche,F.,
             Rodenburg,C.M., Chen,Y., Shaw,D.R., Allen,S., Musonda,R.,
             Shaw,G.M., Zajac,A.J., Letvin,N. and Hahn,B.H.
  TITLE      Codon usage optimization of HIV type 1 subtype C gag, pol, env, and
             nef genes: in vitro expression and immune responses in
             DNA-vaccinated mice
  JOURNAL    AIDS Res. Hum. Retroviruses 19 (9), 817-823 (2003)
REFERENCE    2  (bases 1 to 1485)
  AUTHORS    Gao,F., Li,Y., Decker,J.M., Bibollet-Ruche,F., Rodenburg,C.M.,
             Chen,Y., Shaw,D.R., Allen,S., Musonda,R., Shaw,G.M., Zajac,A.J. and
             Hahn,B.H.
  TITLE      Direct Submission
  JOURNAL    Submitted (16-NOV-2002) Medicine, Duke University Medical Center,
             112 RPIII, LaSalle Street, Durham, NC 27710, USA
FEATURES             Location/Qualifiers
     source          1..1485
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="codon usage optimization of HIV-1 96ZM651"
     gene            1..1485
                     /gene="gag"
     CDS             1..1485

Query Match             100.0%;  Score 60;  DB 6;  Length 1479;
Best Local Similarity   100.0%;  Pred. No. 1.9e-07;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

                     /gene="gag"
                     /codon_start=1
                     /transl_table=11
                     /product="gag protein"
                     /protein_id="AA065393.1"
                     /translation="MGARASILRGGKLDKWEKIRLRPGGKKRYMIKHLVWASRELERF
                     ALNPGLLETSEGCKQIMKQLQPALQTGTEELRSLYNTVATLYCVHEGVERDTKEALD
                     RIEEEQNKIQQKTQQAKAADGKVSQNYPIVQNLQGQMVHQKLSPRTLNAWVKVIEE
                     KAFSPEVIPMFTALSEGATPQDLNTMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVH
                     AGPIAPGQMRBPRGSDIAGTTSTLQEQIAWMTNSNPIPVGDIYKRWIILGLNKIVRMY
                     SPVSILDIKQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTDLLVQNANPDCKTILK
                     ALGPGATLEEMMTACQGVGGPSHKARVLAEAMSQTNSVNILMQKSNFKGNKRVKCFN
                     CGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQNR
                     PEPTAPPAESRFEETTPAPKQESKDREALTSLKSLFGSDPLSQ"
ORIGIN

Qy    1 GACATCAAGCAGGAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
Db  847 GACATCAAGCAGGAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 906

RESULT 8

Qy    1 GACATCAAGCAGGAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
Db  847 GACATCAAGCAGGAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 906

LOCUS        AX455983              1491 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION   Sequence 99 from Patent WO0204493.
ACCESSION    AX455983
VERSION      AX455983.1  GI:21714967
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1
  AUTHORS    zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE      Polynucleotides encoding antigenic hiv type c polypeptides,
             polypeptides and uses thereof
  JOURNAL    Patent: WO 0204493-A 99 17-JAN-2002;
             CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..1491
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Gag_TV2_C_ZAopt"
ORIGIN

Query Match             100.0%;  Score 60;  DB 12;  Length 1491;
Best Local Similarity   100.0%;  Pred. No. 1.9e-07;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

RESULT 9

Qy    1 GACATCAAGCAGGAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
Db  841 GACATCAAGCAGGAGGGGCCCCAAGGAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 900

LOCUS        AX455935              1494 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION   Sequence 51 from Patent WO0204493.
ACCESSION    AX455935
VERSION      AX455935.1  GI:21714919
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1
  AUTHORS    zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,B.
  TITLE      Polynucleotides encoding antigenic hiv type c polypeptides,
             polypeptides and uses thereof

Query Match             100.0%;  Score 60;  DB 6;  Length 1491;
Best Local Similarity   100.0%;  Pred. No. 1.9e-07;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

JOURNAL     Patent: WO 0204493-A 51 17-JAN-2002;
            CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..1494
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="HIV Type C Gag optimized"
ORIGIN

Qy   Query Match          100.0%;  Score 60;  DB 6;  Length 1494;
     Best Local Similarity  100.0%;  Pred. No. 1.9e-07;
     Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

     1 GACATCAAGCACAGGGGCCCCAAGGAGGCCCCTTCCGCGACTACGTGGACCGCCTTCTTCAAGACC 60
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 GACATCAAGCACAGGGGCCCCAAGGAGGCCCCTTCCGCGACTACGTGGACCGCCTTCTTCAAGACC 900

RESULT 10
LOCUS       AX427930                9166 bp    DNA     linear   PAT 20-JUN-2002
AX427930
DEFINITION  Sequence 168 from Patent WO0232943.
ACCESSION   AX427930
VERSION     AX427930.1  GI:21538017
KEYWORDS
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 168 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..9166
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Qy   Query Match          100.0%;  Score 60;  DB 6;  Length 9166;
     Best Local Similarity  100.0%;  Pred. No. 1.5e-07;
     Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

      1 GACATCAAGCACAGGGGCCCCAAGGAGGCCCCTTCCGCGACTACGTGGACCGCCTTCTTCAAGACC 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2723 GACATCAAGCACAGGGGCCCCAAGGAGGCCCCTTCCGCGACTACGTGGACCGCCTTCTTCAAGACC 2782

RESULT 11
LOCUS       AX427931                9169 bp    DNA     linear   PAT 20-JUN-2002
AX427931
DEFINITION  Sequence 169 from Patent WO0232943.
ACCESSION   AX427931
VERSION     AX427931.1  GI:21538018
KEYWORDS
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 169 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..9169
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"

ORIGIN               /note="plasmid pVR1012x/s containing HIV genes"

Qy   Query Match          100.0%;  Score 60;  DB 6;  Length 9169;
     Best Local Similarity  100.0%;  Pred. No. 1.5e-07;
     Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

      1 GACATCAAGCACAGGGGCCCCAAGGAGGCCCCTTCCGCGACTACGTGGACCGCCTTCTTCAAGACC 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2729 GACATCAAGCACAGGGGCCCCAAGGAGGCCCCTTCCGCGACTACGTGGACCGCCTTCTTCAAGACC 2788

RESULT 12
LOCUS       AX427938                9785 bp    DNA     linear   PAT 20-JUN-2002
AX427938
DEFINITION  Sequence 176 from Patent WO0232943.
ACCESSION   AX427938
VERSION     AX427938.1  GI:21538025
KEYWORDS
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 176 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..9785
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"

RESULT 13
LOCUS       AX427936                9788 bp    DNA     linear   PAT 20-JUN-2002
AX427936
DEFINITION  Sequence 174 from Patent WO0232943.
ACCESSION   AX427936
VERSION     AX427936.1  GI:21538023
KEYWORDS
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 174 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..9788
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Qy   Query Match          100.0%;  Score 60;  DB 6;  Length 9785;
     Best Local Similarity  100.0%;  Pred. No. 1.5e-07;
     Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

      1 GACATCAAGCACAGGGGCCCCAAGGAGGCCCCTTCCGCGACTACGTGGACCGCCTTCTTCAAGACC 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2729 GACATCAAGCACAGGGGCCCCAAGGAGGCCCCTTCCGCGACTACGTGGACCGCCTTCTTCAAGACC 2788

Qy        1 GACATCAAGCAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
            ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db     2723 GACATCAAGCAGCAGGGCCCCAAGGAGCCCTTCCCGACTACGTGGACCGCTTCTTCAAGACC 2782

RESULT 14
LOCUS       AX455886                  60 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 2 from Patent WO0204493.
ACCESSION   AX455886
VERSION     AX455886.1  GI:21714879
KEYWORDS
SOURCE      Human immunodeficiency virus
  ORGANISM  Human immunodeficiency virus
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 2 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..60
                     /organism="Human immunodeficiency virus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:12721"
ORIGIN

  Query Match             92.0%;  Score 55.2;  DB 6;  Length 60;
  Best Local Similarity   95.0%;  Pred. No. 7.2e-06;
  Matches    57;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 GACATCAAGCAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
            ||||||   |||||||||||||||||||||||||   |||||||||||||||||||||||
Db        1 GACATCCGCGCCAGGGCCCCAAGGAGCCCTTCCCGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 15
LOCUS       AX455888                1509 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 4 from Patent WO0204493.
ACCESSION   AX455888
VERSION     AX455888.1  GI:21714881
KEYWORDS
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 4 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (2A)
FEATURES             Location/Qualifiers
     source          1..1509
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="synthetic Gag of HIV strain AF110967"
ORIGIN

  Query Match             92.0%;  Score 55.2;  DB 6;  Length 1509;
  Best Local Similarity   95.0%;  Pred. No. 4.8e-06;
  Matches    57;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 GACATCAAGCAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
            ||||||   |||||||||||||||||||||||||   |||||||||||||||||||||||
Db      841 GACATCCGCGCCAGGGCCCCAAGGAGCCCTTCCCGACTACGTGGACCGCTTCTTCAAGACC 900

```
                          GenCore version 5.1.6
               Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on:              May 28, 2004, 07:10:24 ; Search time 27.8049 Seconds
                                              (without alignments)
                                              9167.164 Million cell updates/sec

Title:             US-09-475-704A-1
Perfect score:     60
Sequence:          1 gacatcaagcagggccccaa..........tggaccgcttcttcaagacc 60
Scoring table:     IDENTITY_NUC
                   Gapop 10.0 , Gapext 1.0

Searched:          3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters:      6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing:   Minimum Match 0%
                   Maximum Match 100%
                   Listing first 45 summaries

Database :         N_Geneseq 29Jan04:*
                   1:  geneseqn1980s:*
                   2:  geneseqn1990s:*
                   3:  geneseqn2000s:*
                   4:  geneseqn2001as:*
                   5:  geneseqn2001bs:*
                   6:  geneseqn2002s:*
                   7:  geneseqn2003as:*
                   8:  geneseqn2003bs:*
                   9:  geneseqn2003cs:*
                  10:  geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

## SUMMARIES

| Result No. | Score | % Query Match | Length | DB ID | Description |
|---|---|---|---|---|---|
| 1 | 60 | 100.0 | 60 | 3 | AAA51607 | HIV Gag m |
| 2 | 60 | 100.0 | 1479 | 6 | ABL39974 | Synthetic |
| 3 | 60 | 100.0 | 1479 | 3 | AAA51609 | HIV synth |
| 4 | 60 | 100.0 | 1479 | 9 | AAA51625 | HIV codon |
| 5 | 60 | 100.0 | 1479 | 6 | AAL44548 | HIV-1 p55 |
| 6 | 60 | 100.0 | 1479 | 6 | AAL44552 | HIV-1 p55 |
| 7 | 60 | 100.0 | 1479 | 6 | ABL39953 | Synthetic |
| 8 | 60 | 100.0 | 1479 | 6 | ABL39957 | Synthetic |
| 9 | 60 | 100.0 | 1491 | 6 | ABL40020 | Synthetic |
| 10 | 60 | 100.0 | 1494 | 6 | ABL39972 | Synthetic |
| 11 | 60 | 100.0 | 1494 | 7 | ACA03590 | Synthetic |
| 12 | 60 | 100.0 | 2742 | 7 | ACA03524 | Synthetic |
| 13 | 60 | 100.0 | 2742 | 9 | ADC13278 | DNA of HI |
| 14 | 60 | 100.0 | 2742 | 9 | ADC13241 | DNA of HI |
| 15 | 60 | 100.0 | 3162 | 9 | ADC13239 | DNA of HI |
| 16 | 60 | 100.0 | 3462 | 9 | ADC13237 | DNA of HI |
| 17 | 60 | 100.0 | 3531 | 9 | ADC13234 | DNA of HI |
| 18 | 60 | 100.0 | 3537 | 9 | ADC13236 | DNA of HI |
| 19 | 60 | 100.0 | 3538 | 9 | ADC13235 | DNA of HI |
| 20 | 60 | 100.0 | 3930 | 9 | ADC13230 | DNA of HI |
| 21 | 60 | 100.0 | 3930 | 9 | ADC13231 | DNA of HI |
| 22 | 60 | 100.0 | 3930 | 9 | ADC13232 | DNA of HI |
| 23 | 60 | 100.0 | 4419 | 7 | ACA03523 | Synthetic |

---

```
RESULT 1
AAA51607

ID    AAA51607 standard; DNA; 60 BP.
XX
AC    AAA51607;
XX
DT    31-OCT-2000 (first entry)
XX
DE    HIV Gag major homology region nucleotides 844-903.
XX
KW    Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW    DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS    Human immunodeficiency virus.
XX
XX
PN    WO200003304-A2.
XX
PD    06-JUL-2000.
XX
PF    30-DEC-1999;    99WO-US031273.
XX
PR    31-DEC-1998;    98US-01144995P.
PR    01-SEP-1999;    99US-0152195P.
XX
PA    (CHIR ) CHIRON CORP.
XX
PI    Barnett S, Zur Megede J;
XX
DR    WPI; 2000-452401/39.
XX
PT    Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env
PT    polypeptide and the polypeptide useful for immunizing a mammal especially
PT    human against HIV.
XX
PS    Claim 1; Page 92; 113pp; English.
XX
CC    Expression cassettes comprising a polynucleotide encoding antigenic type
CC    C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful
CC    in DNA immunization, generation of packaging cell lines and production of
CC    Gag- and/or Env-containing proteins. Synthetic Env and Gag expression
CC    cassettes exhibit increased potency for induction of cytotoxic T-
CC    lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-
CC    assemble into non-infectious virus-like particles which are used as a
CC    matrix for the proper presentation of an antigen entrapped or associated
CC    to the immune system of the host
```

## ALIGNMENTS

| | | | | |
|---|---|---|---|---|
| 24 | 60 | 100.0 | 4419 | 9 | ADC13240 | DNA of HI |
| 25 | 60 | 100.0 | 4423 | 9 | ADC13256 | DNA of HI |
| 26 | 60 | 100.0 | 4546 | 9 | ADC13255 | DNA of HI |
| 27 | 60 | 100.0 | 4615 | 9 | ADC13257 | DNA of HI |
| 28 | 60 | 100.0 | 4702 | 9 | ADC13259 | DNA of HI |
| 29 | 60 | 100.0 | 4713 | 7 | ACA03592 | Synthetic |
| 30 | 60 | 100.0 | 4713 | 9 | ADC13280 | DNA of HI |
| 31 | 60 | 100.0 | 4716 | 9 | ACA03522 | Synthetic |
| 32 | 60 | 100.0 | 4716 | 9 | ADC13238 | DNA of HI |
| 33 | 60 | 100.0 | 4716 | 9 | ACA03521 | Synthetic |
| 34 | 60 | 100.0 | 5145 | 7 | ADC13233 | DNA of HI |
| 35 | 60 | 100.0 | 5145 | 7 | ACA03591 | Synthetic |
| 36 | 60 | 100.0 | 5184 | 7 | ADC13279 | DNA of HI |
| 37 | 60 | 100.0 | 5184 | 9 | ABK91616 | Modified |
| 38 | 60 | 100.0 | 9166 | 9 | ABK91616 | Modified |
| 39 | 60 | 100.0 | 9169 | 6 | ABK91617 | Modified |
| 40 | 60 | 100.0 | 9785 | 6 | ABK91624 | Modified |
| 41 | 58.4 | 97.3 | 9788 | 6 | ABK91622 | Modified |
| 42 | 58.4 | 97.3 | 1092 | 7 | Aad53096 | Gag p17/2 |
| 43 | 58.4 | 97.3 | 1092 | 7 | Aad53097 | Gag p17/2 |
| 44 | 58.4 | 97.3 | 3020 | 7 | Aad53105 | Human imm |
| 45 | 58.4 | 97.3 | 3021 | 7 | Aad53106 | Human imm |
| | | | 3021 | 7 | Aad53107 | Human imm |

SQ  Sequence 60 BP; 14 A; 22 C; 15 G; 9 T; 0 U; 0 Other;

Qy  Query Match
    Best Local Similarity  100.0%;   Score 60;   DB 3;   Length 60;
    Matches 60;  Conservative 0;  Pred. No. 9.4e-10;  Indels 0;  Gaps 0;

Db

  1 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  1 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 2
ID  ABL39974 standard; DNA; 60 BP.
XX
AC  ABL39974;
XX
DT  15-MAY-2002 (first entry)
XX
DE  Synthetic Gag polynucleotide sequence SEQ ID NO:53.
XX
DE  Human immunodeficiency virus type C; antigenic HIV type C protein;
KW  immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW  immunostimulant; gene therapy; gene; ds.
XX
OS  Human immunodeficiency virus; type C.
OS  Synthetic.
XX
PN  WO2002004493-A2.
XX
PD  17-JAN-2002.
XX
PF  05-JUL-2001; 2001WO-US021241.
XX
PR  05-JUL-2001; 2000US-0610313.
XX
PA  (CHIR ) CHIRON CORP.
PA  (UYST-) UNIV STELLENBOSCH.
XX
PI  Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg BJ;
XX
DR  WPI; 2002-154920/20.
XX
PT  New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT  applications including DNA immunization or generation of packaging cell
PT  lines, particularly in gene therapy.
PS  Example 1; Fig 24; 233pp; English.
XX
CC  The present invention describes expression cassettes comprising a
CC  polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC  type C polypeptides. The expression cassettes comprise any of the HIV
CC  type C polypeptides sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC  (I). (I) have immunostimulant activity and can be used in gene therapy.
CC  The HIV type C polynucleotides are useful in applications including DNA
CC  immunisation, generation of packaging cell lines, and production of HIV
CC  Type C proteins. The polynucleotides are particularly useful in gene
CC  therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC  ABB06204 to ABB06215 represent sequences used in the exemplification of
CC  the present invention

SQ  Sequence 60 BP; 14 A; 22 C; 15 G; 9 T; 0 U; 0 Other;

Qy  Query Match
    Best Local Similarity  100.0%;   Score 60;   DB 6;   Length 60;
    Matches 60;  Conservative 0;  Pred. No. 9.4e-10;  Indels 0;  Gaps 0;

Db

  1 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  1 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60

RESULT 3

ID  AAA51609 standard; DNA; 1479 BP.
XX
AC  AAA51609;
XX
DT  31-OCT-2000 (first entry)
XX
DE  HIV synthetic Gag polynucleotide.
XX
DB  Gag; expression cassette; antigenic; type C; HIV; Env; synthetic;
KW  DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS  Human immunodeficiency virus; type C strain AF110965.
OS  Synthetic.
XX
FH  Key            Location/Qualifiers
FT  CDS            1..1479
FT                 /tag= a
FT                 /product= "Synthetic_Gag"
FT                 /note= "Codon usage pattern was modified and inhibitory
FT                 elements (INS) and RRE sites were inactivated resulting
FT                 in improved expression"
XX
PN  WO200039304-A2.
XX
PD  06-JUL-2000.
XX
PD  30-DEC-1999;   99WO-US031273.
XX
PR  31-DEC-1998;   98US-0114495P.
PR  01-SEP-1999;   99US-0152195P.
XX
PA  (CHIR ) CHIRON CORP.
XX
PI  Barnett S, Zur Megede J;
XX
DR  WPI; 2000-452401/39.
DR  P-PSDB; AAY96943.
XX
PT  Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env
PT  polypeptide and the polypeptide useful for immunizing a mammal especially
PT  human against HIV.
PS  Claim 2; Page 92-93; 113pp; English.
XX
CC  Expression cassettes comprising a polynucleotide encoding antigenic type
CC  C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful
CC  in DNA immunization, generation of packaging cell lines and production of
CC  Gag- and/or Env-containing proteins. Synthetic Env and Gag expression
CC  cassettes exhibit increased potency for induction of cytotoxic T-
CC  lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-
CC  assemble into non-infectious virus-like particles which are used as a
CC  matrix for the proper presentation of an antigen entrapped or associated
CC  to the immune system of the host

SQ  Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Qy  Query Match
    Best Local Similarity  100.0%;   Score 60;   DB 3;   Length 1479;
    Matches 60;  Conservative 0;  Pred. No. 1.3e-09;  Indels 0;  Gaps 0;

Db

  1 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
844 GACATCAAGCAGGGGCCCCAAGGAGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACC 903

RESULT 4
AAA51625
ID  AAA51625 standard; DNA; 1479 BP.
XX
AC  AAA51625;
XX
DT  31-OCT-2000 (first entry)

XX
DB  HIV codon-optimized synthetic Gag polynucleotide.
XX
KW  Gag; expression cassette; antigenic; type C; HIV; synthetic;
KW  DNA immunization; packaging cell line; antigen presentation; ss.
XX
OS  Human immunodeficiency virus; type C strain AF110965.
OS  Synthetic.
XX
PN  WO200039304-A2.
XX
PD  06-JUL-2000.
XX
PF  30-DEC-1999;  99WO-US031273.
XX
PR  31-DEC-1998;  98US-0114495P.
PR  01-SEP-1999;  99US-0152195P.
XX
PA  (CHIR ) CHIRON CORP.
XX
PI  Barnett S, Zur Megede J;
XX
DR  WPI; 2000-452401/39.
XX
PT  Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV Env
PT  polypeptide and the polypeptide useful for immunizing a mammal especially
PT  human against HIV.
XX
PS  Disclosure; Page 103-104; 113pp; English.
XX
CC  Expression cassettes comprising a polynucleotide encoding antigenic type
CC  C human immunodeficiency virus (HIV) Gag or Env polypeptides are useful
CC  in DNA immunization, generation of packaging cell lines and production of
CC  Gag- and/or Env-containing proteins. Synthetic Env and Gag expression
CC  cassettes exhibit increased potency for induction of cytotoxic T-
CC  lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1 self-
CC  assemble into non-infectious virus-like particles which are used as a
CC  matrix for the proper presentation of an antigen entrapped or associated
CC  to the immune system of the host
XX
SQ  Sequence 1479 BP; 325 A; 533 C; 461 G; 160 T; 0 U; 0 Other;

Query Match           100.0%;  Score 60;  DB 3;  Length 1479;
Best Local Similarity 100.0%;  Pred. No. 1.3e-09;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

```
Qy   1 GACATCAAGCAGGAGGGCCCAAGGAGGAGCCCCTTCCGCGGACTACGTGGACCGGCTTCTTCAAGACC  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 844 GACATCAAGCAGGAGGGCCCAAGGAGGAGCCCCTTCCGCGGACTACGTGGACCGGCTTCTTCAAGACC 903
```

RESULT 5
ID  AAL44548 standard; DNA; 1479 BP.
XX
AC  AAL44548;
XX
DT  29-AUG-2003  (revised)
DT  08-NOV-2002  (first entry)
XX
DE  HIV-1 p55gag polypeptide coding sequence 1.
XX
KW  HIV; ds; vaccine; gene; immune response; microparticle;
KW  adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
KW  polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
KW  submicron emulsion; viral infection; bacterial infection;
KW  parasitic infection; HIV-1 p55gag polypeptide.
XX
OS  Human immunodeficiency virus 1.
XX
PN  WO200226209-A2.
XX
PD  04-APR-2002.

XX
PF  28-SEP-2001;  2001WO-US030540.
XX
PR  28-SEP-2000;  2000US-0236105P.
PR  30-AUG-2001;  2001US-0315905P.
XX
PA  (CHIR ) CHIRON CORP.
XX
PI  O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;
PI  Ulmer J, Dubensky TW;
XX
DR  WPI; 2002-519084/55.
XX
PT  A microparticle to which a biologically active macromolecule is adsorbed,
PT  for use as a vaccine composition to treat viral, bacterial or parasitic
PT  infections, comprises a polymer microparticle, a detergent and a
PT  submicron emulsion.
XX
PS  Claim 72; Fig 1; 100pp; English.
XX
CC  The invention relates to a method of raising an immune response in a host
CC  animal. The method of the invention comprises administering a
CC  microparticle that has an adsorbent surface to which a first biologically
CC  active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC  microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC  acid, a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC  a polycyanoacrylate, a detergent, and submicron emulsion. The method/
CC  microparticle of the invention is useful for immunising a host animal
CC  against viral, bacterial or parasitic infections. The present DNA
CC  sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to
CC  standardise OS field)
XX
SQ  Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 U; 0 Other;

Query Match           100.0%;  Score 60;  DB 6;  Length 1479;
Best Local Similarity 100.0%;  Pred. No. 1.3e-09;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

```
Qy   1 GACATCAAGCAGGAGGGCCCAAGGAGGAGCCCCTTCCGCGGACTACGTGGACCGGCTTCTTCAAGACC  60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 844 GACATCAAGCAGGAGGGCCCAAGGAGGAGCCCCTTCCGCGGACTACGTGGACCGGCTTCTTCAAGACC 903
```

RESULT 6
ID  AAL44552 standard; DNA; 1479 BP.
XX
AC  AAL44552;
XX
DT  29-AUG-2003  (revised)
DT  08-NOV-2002  (first entry)
XX
DE  HIV-1 p55gag polypeptide coding sequence 3.
XX
KW  HIV; ds; vaccine; gene; immune response; microparticle;
KW  adsorbent surface; poly(alpha-hydroxy acid); polyhydroxy butyric acid;
KW  polycaprolactone; polyorthoester; polycyanoacrylate; detergent;
KW  submicron emulsion; viral infection; bacterial infection;
KW  parasitic infection; HIV-1 p55gag polypeptide.
XX
OS  Human immunodeficiency virus 1.
XX

XX
FH  Key              Location/Qualifiers
FT  variation        replace(282, C)
FT                   /*tag= a
FT  variation        replace(552, G)
FT                   /*tag= b
FT  variation        replace(556, T)
FT                   /*tag= c
FT  variation        replace(786, G)
FT                   /*tag= d
FT  variation        replace(819, G)
FT                   /*tag= e

FT   variation       replace(1002, T)
FT   /*tag= f
FT   variation       replace(1005, C)
FT   /*tag= g
FT   variation       replace(1092, G)
FT   /*tag= h
FT   variation       replace(1149, G)
FT   /*tag= i
FT   variation       replace(1158, C)
FT   /*tag= j
XX
PN   WO20022620S-A2.
XX
PD   04-APR-2002.
XX
PF   28-SEP-2001; 2001WO-US030540.
XX
PR   28-SEP-2000; 2000US-0236105P.
PR   30-AUG-2001; 2001US-0315905P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   O'hagan D, Otten G, Donnelly JJ, Polo JM, Barnett S, Singh M;
PI   Ulmer J, Dubensky TW;
XX
DR   WPI; 2002-519084/55.
XX
PT   A microparticle to which a biologically active macromolecule is adsorbed,
PT   for use as a vaccine composition to treat viral, bacterial or parasitic
PT   infections, comprises a polymer microparticle, a detergent and a
PT   submicron emulsion.
XX
PS   Claim 72; Fig 5; 100pp; English.
XX
CC   The invention relates to a method of raising an immune response in a host
CC   animal. The method of the invention comprises administering a
CC   microparticle that has an adsorbent surface to which a first biologically
CC   active macromolecule (e.g. a nucleic acid) has been adsorbed. The
CC   microparticle comprises a polymer microparticle of poly(alpha-hydroxy
CC   acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester,
CC   a polyanhydride, a detergent, and submicron emulsion. The method/
CC   microparticle of the invention is useful for immunising a host animal
CC   against viral, bacterial or parasitic infections. The present DNA
CC   sequence encodes a HIV-1 p55gag polypeptide. (Updated on 29-AUG-2003 to
CC   standardise OS field)
XX
SQ   Sequence 1479 BP; 325 A; 534 C; 460 G; 160 T; 0 U; 0 Other;

Query Match             100.0%;  Score 60;  DB 6;  Length 1479;
Best Local Similarity   100.0%;  Pred. No. 1.3e-09;
Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 GACATCAAGCAGGGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         844 GACATCAAGCAGGGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 903

RESULT 7
ID   ABL39953 standard; DNA; 1479 BP.
XX
AC   ABL39953;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Synthetic Gag polynucleotide sequence SEQ ID NO:3.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.

XX
PN   WO20020493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg BJ;
XX
DR   WPI; 2002-154920/20.
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Example 1; Fig 1; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (II) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 1479 BP; 325 A; 463 C; 462 G; 162 T; 0 U; 0 Other;

Query Match             100.0%;  Score 60;  DB 6;  Length 1479;
Best Local Similarity   100.0%;  Pred. No. 1.3e-09;
Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 GACATCAAGCAGGGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         844 GACATCAAGCAGGGCCCCAAGGAGGAGCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 903

RESULT 8
ID   ABL39957 standard; DNA; 1479 BP.
XX
AC   ABL39957;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Synthetic Gag polynucleotide sequence SEQ ID NO:20.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.

PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2002-154920/20.
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Example 1; Fig 5; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 1479 BP; 325 A; 533 C; 461 G; 160 T; 0 U; 0 Other;

Qy   Query Match     100.0%;  Score 60;  DB 6;  Length 1479;
Best Local Similarity  100.0%;  Pred. No. 1.3e-09;
Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1  GACATCAAGCAGGAGGGGCCCCAAGGAGCGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  844  GACATCAAGCAGGAGGGGCCCCAAGGAGCGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC  903

RESULT 9
ID   ABL40020
AC   ABL40020 standard; DNA; 1491 BP.
XX
DT   15-MAY-2002  (first entry)
XX
DE   Synthetic Gag polynucleotide sequence SEQ ID NO:99.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
PN   WO200204493-A2.
XX
PD   17-JAN-2002.
XX
PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
PA   (UYST-) UNIV STELLENBOSCH.
PA   (CHIR ) CHIRON CORP.
XX
PR   05-JUL-2000; 2000US-0610313.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
DR   WPI; 2002-154920/20.
XX
PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Claim 8; Fig 70; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV

---

CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 1491 BP; 319 A; 557 C; 457 G; 158 T; 0 U; 0 Other;

Qy   Query Match     100.0%;  Score 60;  DB 6;  Length 1491;
Best Local Similarity  100.0%;  Pred. No. 1.3e-09;
Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1  GACATCAAGCAGGAGGGGCCCCAAGGAGCGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  GACATCAAGCAGGAGGGGCCCCAAGGAGCGCCCTTCCGCGACTACGTGGACCGGCTTCTTCAAGACC  900

RESULT 10
ID   ABL39972
AC   ABL39972 standard; DNA; 1494 BP.
XX
DT   15-MAY-2002  (first entry)
XX
DE   Synthetic Gag polynucleotide sequence SEQ ID NO:51.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
PN   WO200204493-A2.
XX
PD   17-JAN-2002.
XX
PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
PA   (UYST-) UNIV STELLENBOSCH.
PA   (CHIR ) CHIRON CORP.
XX
PR   05-JUL-2000; 2000US-0610313.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
DR   WPI; 2002-154920/20.
XX
PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Claim 7; Fig 22; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 1494 BP; 325 A; 557 C; 455 G; 157 T; 0 U; 0 Other;

```
Qy          1 GACATCAAGCAGGGCCCCAAGGAGAGCCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GACATCAAGCAGGGGCCCCAAGGAGAGCCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 900
```

RESULT 11
ACA03590
ID    ACA03590 standard; DNA; 2742 BP.
XX
AC    ACA03590;
XX
DT    22-MAY-2003  (first entry)
XX
DE    Synthetic DNA ecoding immunogenic HIV peptide #73.
XX
KW    Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW    gene therapy; packaging cell line; humoral immune response;
KW    cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS    Synthetic.
XX
PN    WO2003004657-A1.
XX
PD    16-JAN-2003.
XX
PF    05-JUL-2002; 2002WO-US021421.
XX
PR    05-JUL-2001; 2001US-0303192P.
PR    31-AUG-2001; 2001US-0316860P.
PR    16-JAN-2002; 2002US-0349728P.
PR    16-JAN-2002; 2002US-0349793P.
PR    16-JAN-2002; 2002US-0349871P.
XX
PA    (CHIR ) CHIRON CORP.
XX
PI    Zur Megede J,  Barnett SW,  Lian Y;
XX
DR    WPI; 2003-221602/21.
XX
PT    New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT    polypeptides, useful as immunogenic compositions or vaccines for
PT    generating humoral or cellular immune responses against HIV in a subject,
PT    especially humans.
XX
PS    Example 1; Fig 78; 262pp; English.
XX
CC    The invention describes a synthetic polynucleotide encoding 2 or more
CC    immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC    derived from different HIV subtypes. The polynucleotide is useful for
CC    immunisation, generation of packaging cell lines, or production of HIV
CC    polypeptides. The polynucleotide and its encoded proteins are useful as
CC    immunogenic compositions or vaccines for generating humoral or cellular
CC    immune responses against HIV in a subject, or for inducing neutralising
CC    antibodies against HIV. The gene delivery vector comprising the
CC    polynucleotide is also useful for DNA immunisation of, or for generating
CC    an immune response (e.g. a humoral or cellular immune response) in, a
CC    subject such as a mammal, particularly a human. This sequence encodes a
CC    human immunodeficiency virus immunogenic peptide
XX
SQ    Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 U; 0 Other;

Query Match             100.0%;  Score 60;  DB 7;  Length 2742;
Best Local Similarity   100.0%;  Pred. No. 1.4e-09;
Matches    60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATCAAGCAGGGGCCCCAAGGAGAGCCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2104 GACATCAAGCAGGGGCCCCAAGGAGAGCCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 2163
```

RESULT 12
ACA03524
ID    ACA03524 standard; DNA; 2742 BP.
XX
AC    ACA03524;
XX
DT    22-MAY-2003  (first entry)
XX
DE    Synthetic DNA ecoding immunogenic HIV peptide #7.
XX
KW    Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW    gene therapy; packaging cell line; humoral immune response;
KW    cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS    Synthetic.
XX
PN    WO2003004657-A1.
XX
PD    16-JAN-2003.
XX
PF    05-JUL-2002; 2002WO-US021421.
XX
PR    05-JUL-2001; 2001US-0303192P.
PR    31-AUG-2001; 2001US-0316860P.
PR    16-JAN-2002; 2002US-0349728P.
PR    16-JAN-2002; 2002US-0349793P.
PR    16-JAN-2002; 2002US-0349871P.
XX
PA    (CHIR ) CHIRON CORP.
XX
PI    Zur Megede J,  Barnett SW,  Lian Y;
XX
DR    WPI; 2003-221602/21.
XX
PT    New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT    polypeptides, useful as immunogenic compositions or vaccines for
PT    generating humoral or cellular immune responses against HIV in a subject,
PT    especially humans.
XX
PS    Example 1; Fig 12; 262pp; English.
XX
CC    The invention describes a synthetic polynucleotide encoding 2 or more
CC    immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC    derived from different HIV subtypes. The polynucleotide is useful for
CC    immunisation, generation of packaging cell lines, or production of HIV
CC    polypeptides. The polynucleotide and its encoded proteins are useful as
CC    immunogenic compositions or vaccines for generating humoral or cellular
CC    immune responses against HIV in a subject, or for inducing neutralising
CC    antibodies against HIV. The gene delivery vector comprising the
CC    polynucleotide is also useful for DNA immunisation of, or for generating
CC    an immune response (e.g. a humoral or cellular immune response) in, a
CC    subject such as a mammal, particularly a human. This sequence encodes a
CC    human immunodeficiency virus immunogenic peptide
XX
SQ    Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 U; 0 Other;

Query Match             100.0%;  Score 60;  DB 7;  Length 2742;
Best Local Similarity   100.0%;  Pred. No. 1.4e-09;
Matches    60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATCAAGCAGGGGCCCCAAGGAGAGCCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GACATCAAGCAGGGGCCCCAAGGAGAGCCCCTTCCGGCGACTACGTGGACCGGCTTCTTCAAGACC 909
```

RESULT 13
ADC13278
ID    ADC13278 standard; DNA; 2742 BP.
XX
AC    ADC13278;
XX
DT    18-DEC-2003  (first entry)
```

XX
DE   DNA of HIV construct TatRevNefGag_C SEQ ID NO 57.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PI   zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 54; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 U; 0 Other;

Query Match             100.0%;  Score 60;  DB 9;  Length 2742;
Best Local Similarity   100.0%;  Pred. No. 1.4e-09;
Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATCAAGCAGGGGCCCCAAGGAGGGCCCCTTCCGCGACTACGTGGACGGCCTTCTTCAAGACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2104 GACATCAAGCAGGGGCCCCAAGGAGGGCCCCTTCCGCGACTACGTGGACGGCCTTCTTCAAGACC 2163

RESULT 14
ADC13241
ID   ADC13241 standard; DNA; 2742 BP.
XX
AC   ADC13241;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagTatRevNef_C SEQ ID NO 20.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 17; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 2742 BP; 578 A; 1020 C; 859 G; 285 T; 0 U; 0 Other;

Query Match             100.0%;  Score 60;  DB 9;  Length 2742;
Best Local Similarity   100.0%;  Pred. No. 1.4e-09;
Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GACATCAAGCAGGGGCCCCAAGGAGGGCCCCTTCCGCGACTACGTGGACGGCCTTCTTCAAGACC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GACATCAAGCAGGGGCCCCAAGGAGGGCCCCTTCCGCGACTACGTGGACGGCCTTCTTCAAGACC 909

RESULT 15
ADC13239
ID   ADC13239 standard; DNA; 3162 BP.
XX
AC   ADC13239;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagRTmut_C SEQ ID NO 18.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;

```
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, P15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 15; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, P15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3162 BP; 729 A; 1095 C; 975 G; 363 T; 0 U; 0 Other;

Query Match             100.0%;  Score 60;  DB 9;  Length 3162;
Best Local Similarity   100.0%;  Pred. No. 1.4e-09;
Matches   60;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 GACATCAAGCAGGGGCCCCAAGGAGGAGCCCTTCCGGGACTACGTGGACCGGCTTCTTCAAGACC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      850 GACATCAAGCAGGGGCCCCAAGGAGGAGCCCTTCCGGGACTACGTGGACCGGCTTCTTCAAGACC 909
```

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on:  June 1, 2004, 11:29:01 ; Search time 141 Seconds
         (without alignments)
         5939.155 Million cell updates/sec

Title:          US-09-475-704A-4
Perfect score:  1509
Sequence:       1 atgggcgccgccgccagcat..........gcggcccctgagccagtaa 1509
Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched:       682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:   1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued Patents NA:*
            1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
            2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
            3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
            4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
            5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
            6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | % Query Match | Length | DB | Match ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1201 | 79.6 | 1515 | 4 | US-09-475-515-4 | Sequence 4, Appli |
| 2 | 1201 | 79.6 | 4472 | 4 | US-09-475-515-75 | Sequence 75, Appl |
| 3 | 1201 | 79.6 | 4608 | 4 | US-09-475-515-76 | Sequence 76, Appl |
| 4 | 1201 | 79.6 | 4689 | 4 | US-09-475-515-74 | Sequence 74, Appl |
| 5 | 1201 | 79.6 | 4766 | 4 | US-09-475-515-73 | Sequence 73, Appl |
| 6 | 1199 | 79.5 | 2031 | 4 | US-09-475-515-5 | Sequence 5, Appl |
| 7 | 1117.8 | 74.1 | 1853 | 4 | US-09-475-515-5 | Sequence 7, Appl |
| 8 | 1117.8 | 74.1 | 1865 | 4 | US-09-475-515-78 | Sequence 78, Appl |
| 9 | 1117.8 | 74.1 | 1865 | 4 | US-09-475-515-79 | Sequence 79, Appl |
| 10 | 1117.8 | 74.1 | 4319 | 4 | US-09-475-515-6 | Sequence 6, Appl |
| 11 | 1049.2 | 69.5 | 1268 | 4 | US-09-475-515-9 | Sequence 9, Appl |
| 12 | 929.4 | 61.6 | 4307 | 4 | US-09-552-950-2 | Sequence 2, Appl |
| 13 | 925.4 | 61.3 | 9772 | 4 | US-09-552-950-5 | Sequence 5, Appl |
| 14 | 736.6 | 48.8 | 8366 | 4 | US-09-872-733A-6 | Sequence 6, Appl |
| 15 | 734.2 | 48.7 | 1496 | 4 | US-09-184-418C-82 | Sequence 82, Appl |
| 16 | 734.2 | 48.7 | 8972 | 4 | US-09-184-418C-9 | Sequence 9, Appl |
| 17 | 723.8 | 48.0 | 4338 | 4 | US-09-872-733A-1 | Sequence 1, Appl |
| 18 | 709.6 | 47.0 | 1485 | 4 | US-09-184-418C-73 | Sequence 73, Appl |
| 19 | 709.6 | 47.0 | 9010 | 4 | US-09-184-418C-8 | Sequence 8, Appl |
| 20 | 708 | 46.9 | 1476 | 4 | US-09-184-418C-100 | Sequence 100, App |
| 21 | 708 | 46.9 | 8959 | 4 | US-09-184-418C-11 | Sequence 11, Appl |
| 22 | 663.6 | 44.0 | 1486 | 4 | US-09-184-418C-38 | Sequence 38, Appl |
| 23 | 663.6 | 44.0 | 8992 | 4 | US-09-184-418C-4 | Sequence 4, Appli |
| 24 | 651.8 | 43.2 | 8968 | 4 | US-09-184-418C-1 | Sequence 1, Appli |
| 25 | 649.4 | 43.0 | 1479 | 4 | US-09-184-418C-12 | Sequence 12, Appli |
| 26 | 642.2 | 42.6 | 7399 | 2 | US-08-418-848A-9 | Sequence 9, Appli |
| 27 | 642.2 | 42.6 | 9709 | 2 | US-08-188-583-5 | Sequence 5, Appli |
| 28 | 642.2 | 42.6 | 9709 | 3 | US-08-388-353-1 | Sequence 1, Appli |
| 29 | 642.2 | 42.6 | 9709 | 3 | US-08-488-551B-1 | Sequence 1, Appli |
| 30 | 642.2 | 42.6 | 9709 | 4 | US-09-300-572-15 | Sequence 15, Appli |
| 31 | 642.2 | 42.6 | 9709 | 4 | US-09-718-096-15 | Sequence 15, Appli |
| 32 | 642.2 | 42.6 | 12479 | 4 | US-09-318-138-13 | Sequence 13, Appl |
| 33 | 642.2 | 42.6 | 12494 | 3 | US-08-935-312-13 | Sequence 13, Appl |
| 34 | 642.2 | 42.6 | 12494 | 3 | US-08-848-760B-33 | Sequence 33, Appl |
| 35 | 642.2 | 42.6 | 15581 | 3 | US-09-646-538-35 | Sequence 35, Appl |
| 36 | 642.2 | 42.6 | 15581 | 3 | US-09-503-222-35 | Sequence 35, Appl |
| 37 | 633 | 41.9 | 1509 | 4 | US-09-475-515-1 | Sequence 1, Appl |
| 38 | 633 | 41.9 | 1610 | 4 | US-09-020-144-1 | Sequence 1, Appl |
| 39 | 633 | 41.9 | 1845 | 4 | US-09-475-515-2 | Sequence 2, Appli |
| 40 | 633 | 41.9 | 4313 | 4 | US-09-475-515-3 | Sequence 3, Appli |
| 41 | 633 | 41.9 | 9737 | 2 | US-08-944-449-7 | Sequence 7, Appli |
| 42 | 633 | 41.9 | 9737 | 4 | US-09-353-362-7 | Sequence 7, Appli |
| 43 | 632.6 | 41.9 | 2025 | 4 | US-09-475-515-8 | Sequence 8, Appli |
| 44 | 631.4 | 41.8 | 5362 | 3 | US-08-463-210-5 | Sequence 5, Appli |
| 45 | 631.4 | 41.8 | 5362 | 4 | US-08-463-028-5 | Sequence 5, Appli |

## ALIGNMENTS

RESULT 1
US-09-475-515-4
; Sequence 4, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag
US-09-475-515-4

Query Match          79.6%;  Score 1201;  DB 4;  Length 1515;
Best Local Similarity   88.5%;  Pred. No. 9.3e-170;
Matches 1351; Conservative   0; Mismatches 140; Indels  36; Gaps  3;

```
Qy    1 ATGGGCGCCGCCGCCAGCATCCTGCGCGTGGAGAGAAGCTGGAAGCTGGAGAAGAATCCGC  60
          |||||||||||||||||| ||| ||||| ||||| ||||||| |||| |||||||||| |||
Db    7 ATGGGCGCCGCCGCCAGCGTGCTGAGCGGCGGCGAGGAGCTGGACAAGTGGGAGAAGATCCC  66

Qy   61 CTGCGCCCCGGCGGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
          ||||||||||||||||||||||| |  ||| |||| ||  || |||  |||||| ||||||
Db   67 CTGCGCCCCGGCGGGCAAGAAGAAGTACAAGCATCGTGTGGGCCACCGCGCGGAG 126

Qy  121 CTGGAGAGGCTTCGCCCTGAACCCCGGCCGCCTGCTGGAGACACGCCGCGAGGGGCGGATC 180
          ||||||  ||||||||  |||||||||||||||||||||||||  ||| ||||| | |  |
Db  127 CTGGAGCGGCTTCGCCGTGAACCCCGGCCGCCTGCTGGAGACCAGCGGCGAGGAGCTGGCG 186

Qy  181 ATGAAGCAGCTTCCAGCCCGCCCGTCCACCACCGAGGAGAGCTGCGCGCAGCCTGTACAAC 240
          ||   |||||||   |||||||||  |||| ||||||||||||||||| ||||||||||||
Db  187 CTGGGCCAGCTGCAGCCCAGCCCTGCAGACCGGGCAGCGAGGAGAGCTGCGCAGCCTGTACAAC 246
```

RESULT 2
US-09-475-515-75
; Sequence 75, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 75
; LENGTH: 4472
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modUS4.delV1/V2.gag.modSF2
US-09-475-515-75

Query Match     79.6%;  Score 1201;  DB 4;  Length 4472;
Best Local Similarity   88.5%;  Pred. No. 9e-170;
Matches 1351;  Conservative    0;  Mismatches 140;  Indels  36;  Gaps  3;

Qy      1   ATGGGCGCGCCGCCGCCAGCATCCTGCGCGCGGAGAAGCTGGACAAGTGGGAGAAGATCCGC    60
Db   2939   ATGGGCGCGCCGCCCCAGCGGTGCTGGAGCGTCGTTGGGATGGGAGAAGATCCGC       2998

Qy     61   CCCGGCGGGCCACCCTGGAGGAGAGATGATGAGCCGGAG                              120
Db   2999   CTGCCCGCCGCCGCAAGAGAAGTACAAGACTGAAGCACATCGTGTGGGCCCAGCCGCGAG       3058

Qy    121   CTGGAGGGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGGCCTGCAAGCAGATC          180
Db   3059   CTGGAGCGCTTCGCCGTGAACCCCGGCCTGCTGGAGACC                             3118

Qy    181   ATGAAGCAGCTGCAGCCCCTGCAGCCGGCAGCCGAGGAGGTGCGCCTGTACAAC             240
Db   3119   CTGGAGCCAGCTGCAGCCCATGCAGACCCTGCAGCCGGCAGCCGAGCTGCGCCTGTACAAC      3178

Qy    241   ACCGTGGCCACCCTGTACTGCGTGCACCAGCGGATCGAGGTCAAGGACACC                  300
Db   3179   ACCGTGGCCACCCTGTACTGCGTGCACCAGCGGATCGAGGTCAAGGACACC                3238

Qy    301   CTGGAGCAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAAGAGCAG          360
Db   3239   CTGGAGCAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAAGAGCAG         3298

; [DNA/protein sequence alignment continues – Qy (Query) and Db (Database)
; rows with position markers and match bars as shown in the figure]

```
Qy         361 GCCGACGGCA---------------AGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 405
               |||| |||||               |||||||||||||||||||||||||||||||||||
Db        3299 GCCGCCGGCACCGGCAACAGCAGCCAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 3358

Qy         406 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG 465
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3359 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG 3418

Qy         466 ATCGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGC 525
                | |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        3419 GTGGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGC 3478

Qy         526 GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATG 585
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3479 GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATG 3538

Qy         586 CAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTG 645
               |||||||||||||| ||||||||||||||||||||||||||||||||| |||||||||||
Db        3539 CAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTG 3598

Qy         646 CAGGCCGGCCCCGTGGCCCCCGGCCAGATGCGCGACCCCCGCGGCAGCGACATCGCCGGC 705
               || ||||||||| | |||||||||||||||||||| ||||||||||||||||||||||||
Db        3599 CACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC 3658

Qy         706 GCCACCAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCGTGCCCGTG 765
                ||||||||||||||||||||||||||| ||||||||||| |||||||||| | ||||||
Db        3659 ACCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTG 3718

Qy         766 GGCGACATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC 825
               ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3719 GGCGAGATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC 3778

Qy         826 CCCGTGAGCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGAC 885
               |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3779 CCCACCAGCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGAC 3838

Qy         886 CGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACC 945
               ||||||| ||||||||||||||| |||||||||| |||||||||||||||||||||||||
Db        3839 CGCTTCTACAAGACCCTGCGCGCTGAGCAGGCCAGCCAGGACGTGAAGAACTGGATGACC 3898

Qy         946 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGC 1005
               ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        3899 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC 3958

Qy        1006 CCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAC 1065
               ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3959 CCCGCGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAC 4018

Qy        1066 AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGG---CCAACAGCGTGAACATCATGATG 1122
               ||||||||||||||||||||||||||||||||||   | |||   | || ||||||||||
Db        4019 AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGGTGACGAACCCGGCGACCATCATGATG 4078

Qy        1123 CAGAAGAGCAACTTCAAGGGCCCCCGGCGCAACGTCAAGTGCTTCAACTGCGGCAAGGAG 1182
               |||    ||||||||     ||  |||   | ||||||||||||||||||||||||||||
Db        4079 CAGCGCGGCAACTTCCGCAACCAGCGGAAGACCGTCAAGTGCTTCAACTGCGGCAAGGAG 4138

Qy        1183 GGCCACATCGCCAAGAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAG 1242
               ||||||| ||||| ||||||||||||||||||||||||||||||||||   ||||||
Db        4139 GGCCACACCGCCAGGAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGCGCTGCGGCCGC 4198

Qy        1243 GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG 1302
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4199 GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG 4258

Qy        1303 CCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAACCGCAGCGAGCCCGCCGCCCCC 1362
               |||||| |||||||||||||||||||||||||||||| |||                  |
Db        4259 CCCAGCTACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCCG------------------C 4300

Qy        1363 ACCGTGCCCACCGCCCCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGCCCCC 1422
                ||| |||||||||||||||||  ||||||||||||||||  |||   |   || |||||
Db        4301 CCCGAGCCCACCGCCCCCCCCGAGGAGAGCTTCCGCTTCGGCGAGGAGAAGACCACCCCC 4360

Qy        1423 AAGCAGGAGCCCAAGGACCGCGAGCCCTACCGCGAGCCCCTGACCGCCCTGCGCAGCCTG 1482
               |  ||| |||   ||  |  |||     | |   | |||||||||  |||||||||||||
Db        4361 AGCCAGAAGCAGGAGCCCATCGACAAGGAGCTGTACCCCCTGACCAGCCTGCGCAGCCTG 4420

Qy        1483 TTCGGCAGCGGCCCCCTGAGCCAGTAA 1509
               ||||||| || ||||   |||||||||
Db        4421 TTCGGCAACGACCCCAGCAGCCAGTAA 4447
```

```
RESULT 3
US-09-475-515-76
; Sequence 76, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
;  APPLICANT: BARNETT, Susan
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: SRIVASTAVA, Indresh
;  APPLICANT: LIAN, Ying
;  APPLICANT: HARTOG, Karin
;  APPLICANT: LIU, Hong
;  APPLICANT: GREER, Catherine
;  APPLICANT: SELBY, Mark
;  APPLICANT: WALKER, Christopher
;  TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;  TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;  FILE REFERENCE: 1621.002
;  CURRENT APPLICATION NUMBER: US/09/475,515A
;  CURRENT FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 90
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
;   LENGTH: 4608
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: gp160.modSF162.delV2.gag.modSF2
US-09-475-515-76

  Query Match             79.6%;  Score 1201;  DB 4;  Length 4608;
  Best Local Similarity   88.5%;  Pred. No. 9e-170;
  Matches 1351;  Conservative    0;  Mismatches  140;  Indels   36;  Gaps    3;

Qy           1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGAGAAGCTGGACAAGTGGGAGAAGATCCGC 60
               |||||||||||||||||| | ||| ||||||   ||||||||||||||||||||||||||
Db        3075 ATGGGCGCCCGCGCCAGCGTGCTGAGCGGCGGCGAGCTGGACAAGTGGGAGAAGATCCGC 3134

Qy          61 CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG 120
               ||||||||||||||||||||| | |||| |||||||||| | ||||||||||||||||||
Db        3135 CTGCGCCCCGGCGGCAAGAAGAAGTACAAGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 3194

Qy         121 CTGGAGGGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCAAGCAGATC 180
               |||||| |||||||| |||||||||||||||||||||||  ||||||||||   ||||||
Db        3195 CTGGAGCGCTTCGCCGTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCCGCCAGATC 3254

Qy         181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC 240
                ||   ||||||||||||  |||||||||||||| |||||||||||||||||||||||||
Db        3255 CTGGGCCAGCTGCAGCCCAGCCTGCAGACCGGCAGCGAGGAGCTGCGCAGCCTGTACAAC 3314

Qy         241 ACCGTGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTCCGCGACACCAAGGAGGCC 300
               |||||||||||||||||||||||||||    ||||||| |||   |||||||||||||||
Db        3315 ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGACGTCAAGGACACCAAGGAGGCC 3374

Qy         301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAG 360
               ||||| ||||||||||||||||||||||||||| || ||||| |||||||||||   |
Db        3375 CTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGCCGCCGCC 3434

Qy         361 GCCGACGGCA---------------AGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 405
               |||| |||||               |||||||||||||||||||||||||||||||||||
Db        3435 GCCGCCGGCACCGGCAACAGCAGCCAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG 3494

Qy         406 CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG 465
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 4
US-09-475-515-74
; Sequence 74, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
;   APPLICANT: BARNETT, Susan
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: SRIVASTAVA, Indresh
;   APPLICANT: LIAN, Ying
;   APPLICANT: HARTOG, Karin
;   APPLICANT: LIU, Hong
;   APPLICANT: GREER, Catherine
;   APPLICANT: SELBY, Mark
;   APPLICANT: WALKER, Christopher
;   TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;   TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;   FILE REFERENCE: 1621.002
;   CURRENT APPLICATION NUMBER: US/09/475,515A
;   CURRENT FILING DATE: 1999-12-30
;   NUMBER OF SEQ ID NOS: 90
;   SOFTWARE: Patentin Ver. 2.0
US-09-475-515-74
; SEQ ID NO 74
;   LENGTH: 4689
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: gp160.modSF162.gag.modSF2

Query Match   79.6%;   Score 1201;   DB 4;   Length 4689;
Best Local Similarity   88.5%;   Pred. No. 9e-170;
Matches 1351;   Conservative   0;   Mismatches 140;   Indels 36;   Gaps 3;

[DNA sequence alignment, Qy/Db pairs]

; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; FILE REFERENCE: 1621.002
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 4766
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modus4.gag.modSF2
US-09-475-515-73

Query Match     79.6%;  Score 1201;  DB 4;  Length 4766;
Best Local Similarity   88.5%;  Pred. No. 9e-170;
Matches 1351;  Conservative    0;  Mismatches 140;  Indels 36;  Gaps  3;

Qy      1 ATGGGCGCGCCCGCGGCCAGCATCCTGCGCGGCGGCGGAGAAGCTGGACAAGTGGGAGAAGATCCGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3233 ATGGGCGCGCC...GACAAGTGGGAGAAGATCCGC 3292

Qy     61 CTGCGCCCCGGCGGGAAGAAGCACTACATGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 120
Db   3293 CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACATCGTGTGGGCCAGCCGCGAG 3352

Qy    121 CTGGAGCGGCTTCGCCCTGAACCCCGGCGCCTTCAACCCCGACGTGATC 180
Db   3353 CTGGAGCGGCGCTTCGCCGTGAACCCCGGCGCCTTCAACCCCGACGTGATC 3412

Qy    181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC 240
Db   3413 CTGGGACAAGCTGCAGCCCGCCAGCCTTGCAGACCGGCACCGAGGAGCTGCGCTACAAC 3472

Qy    241 ACCGTGGCCACCCTGTACTGCGTGCACGCGCGGCACCCACCAGACGAGGAG 300
Db   3473 ACCGTGGCCACCCTGTACTGCGTCACCACGCGGCACCCAAGGACGAGGCC 3532

Qy    301 CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGACGAGCAGCCAAGGAG 360
Db   3533 CTGGAGAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGGCCCAGCGCCCGCC 3592

Qy    361 GCCGACGGCA------------AGGTGAGCCAGAACTACCCCATCGTGCAGAACGG 405
                        ||||||||||||
Db   3593 GCCGACCGGCGAACCGGCCAACAGCAGCCAGTGAGCAACAGCCAGTGCAGAACCTG 3652

Qy    406 CAGGGCCCAGATGTGGTGCACGCAGGCCATCAGCCCGCGCCCCGGCAGCGAGGTG 465
Db   3653 CAGGGCCCAGATGGTGCACCAGGCCATCAGCCCGCGCCCCGGCAGCGAGGTG 3712

Qy    466 ATCCGAGGAGACGAGGCCTTCAGCCCGAGGTGATCCCCATGTTCACGGCGCGGC 525
Db   3713 GTTGGAGGAGAAGGCCTTCAGCCCGGTGATCCCCATGTTCACGGCCCTG 3772

Qy    526 GCCACCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCG 585
Db   3773 GCCACCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCATG 3832

Qy    586 CAGATGCTGAAGGACACCATCAACGAGGAGGCCGCTGCAGCCCGTG 645
Db   3833 CAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCTGCAGCCCGTGCACCCGTG 3892

Qy    646 CAGGGCCGGCCCCGTGGGCCCCCGGCCAGGACCATGCCGCC 705
Db   3756 CAGATCCTGAAGGACCATCAAACGAGGAGGCCGCCGGCCGCCCCGGTG 3815

Qy    706 GCCACCGACGCACCCTGCAGGAGGAGCAGCAACCCCCCCGTGCCCGTG 765
Db   3876 ACCACCAGCACCCTGCAGGAGGAGCTGGAATGCCAACACCCCCCATCCCGTG 3935

Qy    766 GGCGGACATCTACAAGCGGGTGATCATCGGCGCCTTGAACCAACAGATCGTGCGGATGTACAGC 825
Db   3936 GGCGAGATCTACAAGCGGGTGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC 3995

Qy    826 CCGGTGAGCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGAC 885
Db   3996 CCCACCAGCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGAC 4055

Qy    886 CGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACC 945
Db   4056 CGCTTCTACAAGACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACC 4115

Qy    946 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC 1005
Db   4116 GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC 4175

Qy   1006 CCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCGGGCCAC 1065
Db   4176 CCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCGGGCCAC 4235

Qy   1066 AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGG---CCAACAGCGTGAACATCATGATG 1122
                                                 |||
Db   4236 AAGGCCCGCGTGCTGGCCGAGGCCATGAGCCAGGTGAACAACGCCAACATCATGATG 4295

Qy   1123 CAGAGAGCAACTTCAAGGGCCCCCGGCCAAGACGTCAAGTGCTTCAACTGCGGCAAGGAG 1182
Db   4296 CAGCGCGGCAACTTCCGCAACCAGCGCAAGATCGTGAAGTGCTTCAACTGCGGCAAGGAG 4355

Qy   1183 GGCCACATCGCCAAGAACTGCCGCGGCCCCCCCAAGAAGGGCTGCTGGAAGTGCGGCAAG 1242
Db   4356 GGCCACATCGCCAAGAACTGCCGCGCCCCCCCCAAGAAGGGCTGCTGGAAGTGCGGCAAG 4415

Qy   1243 GAGGGCCACCAGATGAAGGACTGCACCGAGCGGCCAGGCCAACTTCCTGCAGAGCCGG 1302
Db   3473 GAGGGCCACACCGCCAAGGACTGCACCGAGCGGCAGGCCAACTTCCTGGGCAAGATCTGG 3532

Qy   1303 CCCAGCCACAAGGGCCGCCCCGGGCCACTTCCTGCACGAGGCCGGGCCGCC 1362
Db   3533 CTGGAGAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGGCCCAGGCCGCC 3592

Qy   1363 ACCGTGCCCACCGCCCCGCCCCGGAGGAGTTCGGCTTCGGCGAGGAGACCACCCCGCCCC 1422
Db   4518 CCCGAGCCCACCGCGCCCCCCGAGGAGTTCGGCTTCGGCGAGGAGACCACCCCGCCCC 4577

Qy   1423 AAGCAGGAGCCCAAGGACCGGCAGCCCTACCCGCGGCCCTGCGCACCCTG 1482
Db   3653 CAGGGCCCAGATGGTGCACCAGGCCATCAGCCCGCGCCCCGGCAGCGAGGTG 3712

Qy   1483 TTCGGCCAGCGCCCCCTGACCCAGTAA 1509
              || || | ||||||   |||||||||
Db   4638 TTCGGCAACGACCCCAGCAGCCAGTAA 4664

Qy   4578 AGCCAGAAGCAGGAGCCCATCGAGAAGGAGCTGTACCCCCTGACCTCGCTG 4637
Db    466 ATCCGAGGAGACGAGGCCTTCAGCCCGGTGATCCCCATGTTCACGGCGGC 525

Qy      1  ATGGGCGCCCGCGCCAGCATCCTCTCGCGCGGCGGAGAAGCTGACAAGTGGGAGAAGATCCGC     60
Db      7  ATGGGCGCCCGGGCGCCGGCCGCTGCTGAGGCGCGGGCGAGCTGGAGCAAGTGGGAGAAGATCCGC     66

Qy     61  CTGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAG    120
Db     67  CTGCGCCCCGGCGGCAAGAAGAAGTACAAGCTGAAGCACATCGTGTGGGCCAGCCGCGAG    126

Qy    121  CTGGAGGGCTTCGCCCTGAACCCCGGCCTGCTCGAGACCAGCGAGGGCTGCAAGCAGATC    180
Db    127  CTGGAGCGCTTCGCCGTGAACCCCGGCCTGCTCGAGACCAGCGAGGGCTGCCGCCAGATC    186

Qy    181  ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC    240
Db    187  CTGGGCCAGCTGCACCCCAGCCTGCAGACCGGCAGCGAGGAGCTGCGCAGCCTGTACAAC    246

Qy    241  ACCGTGGCCACCCTGTACTGCGTGCACCAGCGGATCGAGGTCAAGGACACCAAGGAGGCC    300
Db    247  ACCGTGGCCACCCTGTACTGCGTGCACCAGCGGATCGAGATCAAGGACACCAAGGAGGCC    306

Qy    301  CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCGCCGCC    360
Db    307  CTGGACAAGATCGAGGAGGAGCAGAACAAGAGCAAGAAGAAGGCCCAGCAGGCTGCCGCC    366

Qy    361  GCCGGCACGGCA----------AGGTGAGCCAGAACTACCCCATCGTGCAGAACCTG    405
Db    367  GCCGGCGGCACCGAGAACCTGTACTTCCAGGGCAACGTGAACCACCGTGCAGAACCTG    426

Qy    406  CAGGGCGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG    465
Db    427  CAGGGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTG    486

Qy    466  ATCGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGC    525
Db    487  GTGGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGC    546

Qy    526  GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATG    585
Db    547  GCCACCCCCCAGGACCTGAACACGATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATG    606

Qy    586  CAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTG    645
Db    607  CAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTG    666

Qy    646  CAGGCCGGCCCCATCGCCCCCGGCCAGATGCGGGAGCCGCGCGGCAGCGACATCGCCGGC    705
Db    667  CACGCCGGCCCCATCGCCCCCGGCCAGATGCGGGAGCCGCGCGGCAGCGACATCGCCGGC    726

Qy    706  GCCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTG    765
Db    727  ACCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTG    786

Qy    766  GGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC    825
Db    826  GGCGACATTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC    885

Qy    826  CCCGTGAGCATCCTGGACATCCGCCAGGGCCCCAAAGAGCCCTTCCGCGACTACGTGGAC    885
Db   4073  CCCACCAGCATCCTGGACATCCGCCAGGGCCCCAAAGAGCCCTTCCGCGACTACGTGGAC    4132

Qy    886  CGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCGACCCAGGAGGTGAAGAACTGGATGACC    945
Db   4133  CGCTTCTACAAGACCCTGCGCGCCGAGCAGGCGAGCCAGGAGGTGAAGAACTGGATGACC    4192

Qy    946  GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC    1005
Db   4193  GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCTCTCGGC    4252

Qy   1006  CCCGCGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAC    1065
Db   4253  CCCGCGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAC    4312

Qy   1066  AAGGCCCGCGTGCTGGGCGAAGCCCGCGTGAGCCAGG---CCAACAGCGTGAACATCATGATG    1122
Db   4313  AAGGCCCGCGTGCTGGGCGAAGCCCGCGTGAGCCAGGGTGAGCAACATCATGATG    4372

Qy   1123  CAGAGAGGCAACTTCAAGGGCCCCGGCGGCAAGTCAAGTGCTTCAACTGCGGCAAGGAG    1182
Db   4373  CAGCGCGGCAACTTCCGCGAACCACACGGGCAAGTCAAGTGCTTCAACTGCGGCAAGGAG    4432

Qy   1183  GGCCACATCGCCAAGAACTGCCGCGCCCCGCGCAAGAAGGGCTGCTGGCGCTGCGGCAAG    1242
Db   4433  GGCCACACCGCCAAGAACTGCCGCGCCCCGCGCAAGAAGGGCTGCTGGCGCTGCGGCAAG    4492

Qy   1243  GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG    1302
Db   4493  GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG    4552

Qy   1303  CCCAGCCACAAGGGCGGCCCCGGCCACTTCCTGCAGAACCGGCCCGAGCCCACCGCCCCC    1362
Db   4553  CCCAGCTACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCCG-------------C    4594

Qy   1363  ACCGTGCCCACGGCCCCACCGGCCTCCGGCTTCGCGGCCCCGAGCGAGACCACCCCCGCC    1422
Db   4595  CCCGAGCCCACACCGGCCCCACCGGCCTTCGGCGCCCCGAGCGAGACCACCCCCGCC    4654

Qy   1423  AAGCAGGAGAGCCCAAGGACGCCCAAGGCCCCTTACCGGGAGGCCCGAGCCGGAGCCTG    1482
Db   4655  AGCCAGAAGCCAGAGAGCCCCAAGGACGCCTGTACCCCTGACCCAGCCCTGCCCGGAGCCTG    4714

Qy   1483  TTCGGCAGCGGCCCCCTGAGCCAGCAGTAA    1509
Db   4715  TTCGGCAACGACCCCAGCCAGCAGTAA     4741

```
;   NUMBER OF SEQ ID NOS: 90
;   SOFTWARE: Patentin Ver. 2.0
;
;   SEQ ID NO 5
;   LENGTH: 1853
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: HIV-Gag-protease
US-09-475-515-5

Query Match     74.1%;   Score 1117.8;  DB 4;  Length 1853;
Best Local Similarity   85.1%;   Pred. No. 1.8e-157;
Matches 1299;   Conservative  0;  Mismatches 192;  Indels 36;  Gaps 3;
```

```
Qy    1  ATGGGCGCGCCGCCCCCGAGCAATCCTGCCGCCCGGCGGACAAGCTGTGGACAAGATCCGC   60
         |||||  ||| || || |  ||  |||| | || |||  ||||||| |||||| ||||||||
Db    7  ATGGGAGCGGCCGCCGGCGCCAGCCGTGCTGAGCGGGCGGCCGAGCTGGACAAGTGGGAGAAGATCCGC   66

Qy   61  CTTCGGCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTTGTGTGGGCCAGCCGCGAG  120
         ||| || |||||||| ||||||| ||| || |||||||  || | | || |||||||||
Db   67  CTTCGGCCCGGCGGCAAGAAGAAGTACAAGCTGAAGCACATCGTGTGGGCCAGCCGCGAG  126

Qy  121  CTGGAGCGGTTCGCCGTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC  180
         |||||||| |||||| |||||||||||||| ||||| ||||  |||||||||||||||||
Db  127  CTGGAGCGGTTCGCCGTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATC  186

Qy  181  ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAAC  240
         ||||||||||||||||| || ||||||||||||||||| ||||| ||||||||||||||
Db  187  CTGGGCCAGCTGCAGCCCGCCCTGCAGACCGGCAGCGAGGAGCTGCGCAGCCTGTACAAC  246

Qy  241  ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTGAAGGACACCAAGGAGGCC  300
         ||||||||||||||||||||||||| || |||||||| ||| | || ||||||||||||
Db  247  ACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTCAAGGACACCAAGGAGGCC  306

Qy  301  CTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCGCCGAG  360
         ||||| || ||||||||| || || |||| |||||||||||||||||||||||| ||||
Db  307  CTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGCCGCCGCC  366

Qy  361  GCGGACGGCA----------------------AGGTGAGCCAGAACTACCCCATCGTGCAG  405
         ||| ||  ||                      ||||||||||||| |||||| ||||||
Db  367  GCCCGCCGGGCGACGCGGCACAGCAGCAGGTGCAGGTGAGCCAGAACTACCCCATCGTGCAG  426

Qy  406  AACGCCCAGATGGTGCACCAGGCCATCAGCCCCAGCACCCTGAACGCCTGGGTGAAGGTG  465
         || ||||||||||||||||||||||| ||||||||||| ||||| |||||||||||||||
Db  427  AACGCCCAGATGGTGCACCAGGCCATCAGCCCCAGCACCCTGAACGCCTGGGTGAAGGTG  486

Qy  466  ATCGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGC  525
         |||||||| ||||||||||||| |||||||||||||||||| ||||||||||||||||||
Db  487  GTGGAGGAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCAGCGCCCTGAGCGAGGGC  546

Qy  526  GCCACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATG  585
         |||||||||||||||||||||||||||||| ||||| |||||| |||||||||||| |||
Db  547  GCCACCCCCCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATG  606

Qy  586  CAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGACTGGGACCGCCTGCACCCCGTG  645
         |||| ||||||||| |||||||||||||| |||||||| ||||||||||||||||||| |
Db  607  CAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTG  666

Qy  646  CAGGCCGGCCCCGTGGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC  705
         ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db  667  CACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGC  726

Qy  706  GCCACCAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAACAACCCCCCCGTCCCCGTG  765
         ||||||||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db  727  ACCACCAGCACCCTGCAGGAGCAGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTG  786

Qy  766  GGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC  825
         ||||| |||||||||||||| || ||||||||| |||||||||||| ||||| ||||||
Db  787  GGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC  846

Qy  826  CCCGTGAGCGCATCCTGGACCATCCGCCAGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGAC  885
         |||                            |||||||||||||||||||||||||||||||
Db                                                                    885
```

```
Qy  766  GGGGCACATCTACAAGCGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGC  825
         |||                                                          |||
Db  787  GGGGGAGATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGACC  846

Qy  826  CCCGTGAGCGCATCCTGGACCATCCGCCAGGGGCCCCAAGGAGCCCTTCCGGGACTACGTGGAC  885
Db  847  CCCAACACCATCCGGACATCGCCGGCACCACCAGCACCCTGCAGGAGCAGATCGGGGCATGGAC  906

Qy  886  CGCTTCTACAAGACCCTGCGCGCTGAGCCAGGCAGGCCACCCAGGGACGCGGGATGAAGAGATGAACGAGGATGACC  945
Db  907  CGCTTCTACAAGACCCTGCGCGCGGAGCAGGCCAGCCAGGACGTGAAGAACTGGATGACC  966

Qy  946  GAGACCCTGCTGGTGCAGAACGCGCCAACCCCGACGTGCAAGACTGCAAGACCATCCTGCGCGGGTCTCGGC  1005
Db  967  GAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGGCCTCTCGGC  1026

Qy 1006  CCCGGCGCCCACCCTGGAGGAGATGATGACCGCGTGCCTGCAGGGCGGCGCCGGCCGGCCGGCCAC  1065
Db 1027  CCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAC  1086

Qy 1066  AAGGCCCGGGTGCTGGCCGAGGCGATGAGCCAGGTGACGAACAGCG---CCAACATCATGATG  1122
Db 1087  AAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAGGTGACCAACAGCGCCACCATCATGATG  1146

Qy 1123  CAGAAGAGCAACTTCAAGGGCCCCCGGCGCAAGGTCAAGTGCTTCAACTGCGGCAAGGAG  1182
Db 1147  CAGCGCGGCAACTTCCGCGGAACAGCCGTGCTCAAGTGCTTCAACTGCGGCAAGGAG  1206

Qy 1183  GGCCACATCGCCAGGAACTGCCGCGCGCCCCGGAAGAAGGGCTGCTGGAAGTGCGGCAAG  1242
Db 1207  GGCCACACCGCCAGGAACTGCCGCGCGCCCCGGAAGAAGGGCTGCTGGAAGTGCGGCGGC  1266

Qy 1243  GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGG  1302
Db 1267  GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCCAAGATCTGG  1326

Qy 1303  CCCAGCCACAAGGGCCGGCCCGGGCCAACTTCCTGCAGAACCGGCCCGAGCCCACCGCCCCC  1362
Db 1327  CCCAGCTACAAGGGGCGCCCGGGCCAACTTCCTGCAGAGC  1368

Qy 1363  ACCGTGCCCACCGCCCCGCCCCCCGCCGAGAGCTTCCGGCTTCGGGCCGAAGAGACCACCCCGCCCCC  1422
Db 1369  CCCGGAGCCCCACCGCCCCCCGAGAGCTTCGGCTTCGGGGAAGATCACCACCCCC  1428

Qy 1423  AAGCAGGAGCCCAAGGACAGCGGGAGGGCCCTACCGCGAGGCCCCTGGACCGCCGCGCCGCCTG  1482
Db 1429  AGCCAGAAGCAGGAGCCCATCGACAAGGAGCTGTACCCCCTGACCAGCCTGCGCAGCCTG  1488

Qy 1483  TTCGGCAACGACCCCAGCTGAGCCAGT  1507
Db 1489  TTCGGCAACGACCCCAGCTGAGCCAGT  1513
```

```
RESULT 7
US-09-475-515-5
Sequence 5, Application US/0947551SA
Patent No. 6602705
GENERAL INFORMATION:
   APPLICANT: BARNETT, Susan
   APPLICANT: ZUR MEGEDE, Jan
   APPLICANT: SRIVASTAVA, Indresh
   APPLICANT: LIAN, Ying
   APPLICANT: HARTOG, Karin
   APPLICANT: LIU, Hong
   APPLICANT: GREER, Catherine
   APPLICANT: SELBY, Mark
   APPLICANT: WALKER, Christopher
   TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
   TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
   FILE REFERENCE: 1621.002
   CURRENT APPLICATION NUMBER: US/09/475,515A
   CURRENT FILING DATE: 1999-12-30
```

RESULT 8
US-09-475-515-78
; Sequence 78, Application US/09475515A
; Patent No. 602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 78
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GP1
US-09-475-515-78

Query Match             74.1%;  Score 1117.8;  DB 4;  Length 1865;
Best Local Similarity   85.1%;  Pred. No. 1.8e-157;
Matches 1299;  Conservative   0;  Mismatches 192;  Indels  36;  Gaps  3;

[Sequence alignment of Query (Qy) against Database (Db) sequence — DNA base alignment spanning positions 1–1515 (Qy) / 13–1005 and beyond]

RESULT 9
US-09-475-515-79
Sequence 79, Application US/09475515A
Patent No. 6602705

GENERAL INFORMATION:

APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher

TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES

FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 79
LENGTH: 1865
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: GP2

US-09-475-515-79

Query Match      74.1%;    Score 1117.8;   DB 4;   Length 1865;
Best Local Similarity    85.1%;   Pred. No. 1.8e-157;
Matches 1299;  Conservative    0;  Mismatches 192;  Indels 36;  Gaps 3;

RESULT 10
US-09-475-515-6

; Sequence 6, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 4319
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag-polymerase
US-09-475-515-6

Query Match 74.1%; Score 1117.8; DB 4; Length 4319;
Best Local Similarity 85.1%; Pred. No. 1.8e-157;
Matches 1299; Conservative 0; Mismatches 192; Indels 36; Gaps 3;

RESULT 11

US-09-475-515-9

Sequence 9, Application US/09475515A

Patent No. 6602705

GENERAL INFORMATION:
  APPLICANT: BARNETT, Susan
  APPLICANT: ZUR MEGEDE, Jan
  APPLICANT: SRIVASTAVA, Indresh
  APPLICANT: LIAN, Ying
  APPLICANT: HARTOG, Karin
  APPLICANT: LIU, Hong
  APPLICANT: GREER, Catherine
  APPLICANT: SELBY, Mark
  APPLICANT: WALKER, Christopher
  TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
  TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
  FILE REFERENCE: 1621.002
  CURRENT APPLICATION NUMBER: US/09/475,515A
  CURRENT FILING DATE: 1999-12-30
  NUMBER OF SEQ ID NOS: 90
  SOFTWARE: Patentin Ver. 2.0
  SEQ ID NO 9
  LENGTH: 1268
  TYPE: DNA
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: Synthetic Gag
  OTHER INFORMATION: common region

US-09-475-515-9

Query Match 69.5%;    Score 1049.2;    DB 4;    Length 1268;
Best Local Similarity 90.8%;    Pred. No. 2.5e-147;
Matches 1146; Conservative 0; Mismatches 98; Indels 18; Gaps 2;

RESULT 12

Query Match　61.6%;　Score 929.4;　DB 4;　Length 4307;
Best Local Similarity　77.7%;　Pred. No. 1.2e-129;
Matches 1173;　Conservative　0;　Mismatches 306;　Indels 30;　Gaps 3;

```
Qy        1 ATGGGCGCCCCGCGCCAGCCATCCTGCGCGGCGGGAGAAGCTGGACAAGTGGGAGGAAGATCCGC 60
            |||||||||||  |||  || |||
Db        1 ATGGGGCGCCCGCGCCAGCCATCCAGC ...                              60

Qy       61 CTGCGCCCCCGCGGGAAGAGCACTACATTGCTGTGGGGGGGGGCAGCGCCGAG 120
Db       61 CTGCGCCCCCGCGGGAAAAGAAGTACAAGCATCGTGAAGCACATCCGCGGCGAA 120

Qy      121 CTGGGAGGAGCCTTCGCCCTGAACCCCGGCCTGCTGGGAGACCCGGCCTGCAAGCAGATC 180
Db      121 CTGGGAGGACGCTTCGCCCTGAACCCCGGCCTCTCGGAGACCCGGGGTGCCCGCAGATC 180

Qy      181 ATGAAGCAGCTCCAGCCCGCGCCCTGCAGGACCGGGCACCGGGAGGAGCTGGCGCAGCCGTACAAC 240
            |||||||||    ||||
Db      181 CTCGGCCAACTCGAGCCCAGCCTGAAACCCGCCGGAGGAGCTGCCCGAGCCTGTACAAC 240

Qy      241 ACCGGTGGCCACCCTGTACTGCGGCCATCGAAATCAAGGATAACCCGGCCGGAAAGAGGCC 300
Db      241 ACCGGTGGCCACCCTGTACTCGCAGCCGCCGCCATCGGAGGAGCCTGTGCAAAC 300

Qy      301 CTGGGACAAGATCGAGGAGCAGCGACAAGTCCCAGCGAGCAGCCCAAGGAG 360
Db      301 CTGGGATAAATCGAAGAGGAACAGAATAAGGCAAAAAGACCCAACAGCCGCG 360

Qy      361 GCCGACGG--------CAAGGTGAGCCAGAACTACCCCATGCTGCAGAACCTCCAGGG 411
            |||||||
Db      361 GACACCGGACAGCAACCAGGTCAGCCAGAACTACCCCATGCTGCAGAAACATCCAGGG 420

Qy      412 CAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCGTGGGTGAAGGTGATCGAG 471
Db      421 CAGATGGTGCACCAGGCCATCTCCCCCGCACCCTGGGTGAAGGTGGTCGAA 480

Qy      472 GAGAAGGCCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACC 531
Db      481 GAGAAGGCCTTTTAGCCCGGAGAGTGATACCCATGTTCTCAGCGCCCTTGTCAGAGGGCGCCACC 540

Qy      532 CCCCAGGACCTGAACACGATGTTGAACACGTGGGGCGCCAGGCGCCCCATGCAGATG 591
Db      541 CCCCAAGATTCTGAACACCATGCTCAACACAGTGGGCGCCGGACCACCCATGCAGATG 600

Qy      592 CTGAAGGACACCATCAACGAGGAGGCCGCGCGAGTGGACCGCGCCTGCACCCCGTGCAGGCC 651
Db      601 CTGAAGGAGACCATCAATGAGGAGGCTCCCGAATGGGATCGTGCATCCGGTTCCACGCA 660

Qy      652 GGGCCCCGTGGCCCCCGCCGGCCAGATGCGAGAGCCACGGGCCTCAGACATCGCCGGCGAAACC 711
Db      661 GGGCCCCATCGCACGCGGCCCGCCAGATGAGAGCCACGGGCCTCAGACATCGCCGGGAACGACT 720
```

```
Qy      712 AGCACCCTGCCAGAGCAGCAGATCGGCTGGATGACCAACAACCCCGTGCCGGTGGGCGAC 771
Db      721 AGTACCCTTCAGGAACAGATCGGCTGGATGACCAACAACCCACCCATCCCGGTGGGGGAGA 780

Qy      772 ATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCTACC 831
Db      781 ATCTACAAACGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGGATGTATAGCCCTACC 840

Qy      832 AGCATCCTGGACATCCGACAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTC 891
Db      841 AGCATCCTGGACATCCGCCAAGGCCCCAAGGAACCCTTTCGCGACTACGTGGACCGGTTC 900

Qy      892 TTCAAGACCCTGCGCGCCGAGCAGGCCCACCAGGAGGTGAAGAACTGGATGACCGAGACC 951
Db      901 TACAAAAGCCTCCCGCGCCGAGCAGGCCCCAGAGGTGAAGAACTGGATGACCGAAACC 960

Qy      952 CTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGC 1011
Db      961 CTGCTGGTCCAGAACGCCAACCCGGACTGCAAGACGATCCTGAAGGCCCTGGGCCCAGCG 1020

Qy     1012 GCCACCCTGGAGGAGATGAATGACCGGCCTGCCGAGGCGGTGGGGCGGCGCCCAGCGGCA 1071
Db     1021 GCTACCCTTAGAGGAAATGATGACCGGCCTTGTCAGGGGCGTGGGCGGACCCCAGCCACAAGGCA 1080

Qy     1072 CGGCCGTGCTGGCCGAGGCGGTGCAGGGCGGCGG---CCAACAGCGGTGGCCAGAAG 1128
Db     1081 CGGCTGCCTGCTGGGCCTGAGGCGATGAGCCCAGG ... 1140

Qy     1129 AGCAACTTCAAGGGCCCCCGGCGCAAGTCAAGTGCTTCAACTGGGCGCCAAGGAGGCCAC 1188
Db     1141 GGCAACTTCGGGAACCAACGCAAGATCGTCAAGTGCTTCAACTGTGGCAAAGAAGGGCAC 1200

Qy     1189 ATCCCCAAGAACTGCCGCGCCCCGCGCCCCCGCGAAGAAGGGCTGTGGAAGTGGGCAAGGAGGGC 1248
Db     1201 ACAGCCCGCAACTGCAGGGCCCCTAGGAAAAAGGGCTGCTGGAAATGCGGCAAAGAGGGC 1260

Qy     1249 CACCAGATGAAGGACTGCACCGAGCGCCAGGCGAACTTCCTGGCAAGATCTGGCCCAGC 1308
Db     1261 CACCAGATGAAAGACTGTACGGAGCAGACAGGCTAATTTTTTAGGGGAAGATCGGCCTTCC 1320

Qy     1309 CACAAGGGCGGCGCCGGAACTTCCTGCAGAAGAACCGCAGCGGGAGCCCGCCCACCGTG 1368
Db     1321 TACAAGGGAAGGCCAGGGGAATTTTCTTCAGACAG------ACCCGCG 1362

Qy     1369 CCCACCGCCCCGCCCCCCGCCGGAGAGCTTCGCGTTCGGAGGAGACACCCCGCCCCAAGCAG 1428
Db     1363 CCAACAGCCCCACCAGCCCCGCCGGAGCCTTCAGGTTCGGGGTCGGGACAACACTCCCCTCCG 1422

Qy     1429 GAGCCCAAGGACCCCGGAGCCCTACCGGCCCCTCGACCGCCCAGCCTGTTCGGC 1488
Db     1423 AAGCAGGAGCCGGACCCGCCAGGGCACGGGTGCCCTCAAACTTCCCCTCAGGGTCGGC 1482

Qy     1489 AGCGGCCCCC 1497
            |||||||
Db     1483 AGCGACCCC 1491
```

RESULT 13

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:pSYNGP
US-09-552-950-5

Query Match          61.3%;  Score 925.4;  DB 4;  Length 9772;
Best Local Similarity  77.3%;  Pred. No. 4.6e-129;
Matches 1175;  Conservative  0;  Mismatches 316;  Indels 30;  Gaps 3;

```
Qy      1 ATGGGCGCGAGAGCGTCCTCGCGCGCGGGAGAGAAGGATGATGATAAATGGCAGAGGATCCGC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   1108 ATGGGCGCGCCCGCGGGCCCAGCGTCGTGTCGGGGGCGCGGGAGGAAGACTGGC 1167

Qy     61 CTGCGCCCCGGCGGCCGCAAGGAGCACTACAATGCTGAAGGCGCGAGCTGGGAAGAAGCTAC 120
Db   1168 CTGCGCCCCGGCGGCCGCAAAAAGAAGTACAAGCTGAAGCACATCGTGTGGGCCAGCCGC 1227

Qy    121 CTGGGAGGGCTTCGCCCCTGACACCCGGCGGCCTGCTGGAGGACGCCGCAAGCAGATC 180
Db   1228 CTGGGAGGGCTTCGCCCGTGAACCCTGGGCATGTGGGCTGCCCCCAGATC 1287

Qy    241 ACCGTGGCCACCCGTACTGCGTCACGCCGGAGGTCCGCGACCAAGAGGGCC 300
Db   1348 ACCGTGGCCACGGCGTACTGCGTCACGCCGGAGGTCCGCGAA 1407

Qy    301 CTGGCAAGATCGGAGGAGGAGAGAACAAGTCCCAGCAGACGAGACCCAAGAG 360
Db   1408 CTGGATAAATCGAACGAGGATACAGAATAAGAGCAAAAGACCCAACGAGCCG 1467

Qy    361 GCCGACGG---------CAAGGTGAGCCAGAACTACCCGATTGTGCAGAATA 411
Db   1468 GACACCGGACAACAGCAGCCAGTCACCAGAACTACCCCATGTGCAGAAAT 1527

Qy    412 CAGATGGTGCACCAGGCCATCAGCCCGGAGGCCACTCCCCGCACCCGCACC 471
Db   1528 CAGATGGTGCACCAGGCCATCTCCCCCGGAGGCCGCTGGGGTGGAA 1587

Qy    472 GAGAGAAGGGCCTTCAGCCCCGGAGGGTGATCCCCATGTACACGCCCCTGA 531
Db   1588 GAGAGAGGCTTTTAGCCCGGAGGTGATACCCATGTTCTCAGCCCTG 1647

Qy    532 CCCCAGGACCTGACACGATGTTGAACACCGTGGCGGCCCGGCCGCCATGCA 591
Db   1648 CCCCAAGATCTGAACACCATGCTCAACACAGTGGGGGGACACCCGGCCATG 1707

Qy    592 CTGAAGGACACCATCAACGAGGAGGCCGGAGTGGAGGGCCCAAGCAGGCC 651
Db   1708 CTGAAGGAGACCATCAATGAGGAGGCCGGATGTCGTGTCAACCGTGCAGCGA 1767

Qy    652 GGCCCCGTGCGCCCAGATGCCGACCCCGGCGGCGACCATCGCCGGCGGCCACC 711
Db   1768 GGGCGCCATCGCACCGGGCCAGATGCGTGAGCCCAGCAACTGGCCACGACT 1827

Qy    712 AGCACCTTCAGGAGCAGCGACATCGCCGCCTTGGACATCCAGGAAGCCCCCC 771
Db   1828 AGTACCCTTCAGGACGAACGATCGGGCTTGGATGACCAACCCCCATCCCCGTCTGGGGCGAC 1887

Qy    772 ATCTACAAGAGGGTGGATCATCCTGGGGCTGAACAAGATCGTGCGGATGTA 831
Db   1888 ATCTACAAACGCTGGATCATCCTGGGCTGAACACAAGATCGTGCGCATGTA 1947

Qy    832 AGCATCCTTGGACATCCGGCGCAGGGCCCCTTCCCGCGACTACGTGCGACCCCTTC 891
Db   1948 AGCATCCTGGACATCCGCCAAGGGCCGAAGGAAACCCTTCGCGACTACGTGCGGGTTC 2007

Qy    892 TTCAAGACCCTGCGCGCCCCAGCGACGTGAAGAACTGGATGACCGAGACC 951
Db   2008 TACAAAAACGCTCCGGCGCCGAGCAGCCGAGCGTGAAGAACTGGATGACCGAGAACC 2067
```

RESULT 14
US-09-872-733A-6
; Sequence 6, Application US/09872733A
; Patent No. 6656706
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733A
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 6
; LENGTH: 8366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of the construct pCMVgagpolBNhan containing a CMV
; OTHER INFORMATION: promoter, a HIV gag/pol gene and a kanamycin
; OTHER INFORMATION: resistance gene
US-09-872-733A-6

Query Match          48.8%;  Score 736.6;  DB 4;  Length 8366;
Best Local Similarity  69.5%;  Pred. No. 3.6e-101;
Matches 1057;  Conservative  0;  Mismatches 434;  Indels 30;  Gaps 3;

```
Qy    952 CTGCTGTGGGTTGCAGAACGCCAACCCCAGACTTCCAAGAACCATCCTTGCCGCGCTCTCGGGCCCCGGC 1011
Db   2068 CTGCTGGTCCAGAACGCCGAACCCCGACTGCGAAGACGATCCTGAAGGCCTGGGCCCACGG 2127

Qy   1012 GCCACCCTGGAGCGAGATGAGTGACCGCGGAACCCCGGACTGGCAGCAACGGTGACCATCATGATGTGCAGAAG 1071
Db   2128 GCTACCCTTAGAGGAAATGAGTGAGCCGCCTCGTCAGGAGGTGGGGGCCACCAGGGCA 2187

Qy   1072 CGGCGTGCTGGCCCGAGGCGGATGAGCCAGG---CCAACAGCGGCTGAACATCATGATGTGCAGAAG 1128
Db   2188 CGCGTCCGTGGCGTGAGGCCCATGAGCCAGTGACCAAAGTCCGCGACTGCCGCACAAGCGC 2247

Qy   1129 AGCAAACTTCAAGGGCCCCGGGCGCAACGTCAAGTGCTTCAAACTCGCGTACCATCATCGTCAGGTC 1188
Db   2248 GGCAACTTCGGAACCAACGCCAAGATCGTCAAAGTGCCTTCAAACTGTGTGCAAGGAAGGGCAC 2307

Qy   1189 ATCGGCCAAGAACTGCCGGCCGCGCGCAACGTCAAGAAGCGGCTGGTGCTGGGCGCCGGCAAGAAGGGGC 1248
Db   2308 ACAGCCCCGGAACTGCCAGGGCCCCTTAGGGAAAAGGCGCTGTTGGAAATGTGGAAAAGGAGGA 2367

Qy   1249 CACCAGATGAAGGACTGCACCGAGCGGCAGGCCAACCTTCCTGGGGCAAGATCTGGCCAGC 1308
Db   2368 CACCAAATGAAAGATTGTACTGAGAGACAGGCTAATTTTTTAGGGAAGAATCTGGCCTTCC 2427

Qy   1309 GAGCCCCCAAGGACCGCGCGAGCCCTACCCGGCAAGGACCGAGAACGACGCAGGGACCGGAGGCCCCGTG 1368
Db   2368 CCAACAGCGCCCACCAGATGCGAAGAGCTTCAGGTTTGGGAGGAACAACTCCCTCTCAG 2529

Qy   1369 CCCACCGCGCCCCCGGAGAGTTCGGCTTCGGAGGAGGACCACACCCCCCAAGCAG 1428
Db   2470 CCAACAGCGCCACCAGATGCGAAGAGCTTCAGGTTTGGGAGGAACAACTCCCTCTCAG 2529

Qy   1429 GAGCCCAAGGACCGGCGGAGCCCTACCCGCGGCAAGACAACGGCAGGGACCGGTG 1488
Db   2530 AAGCAGGAGCCGGGATAGACAAGGAACTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGC 2589

Qy   1489 AGGGGCCCCCCTGAGCCAGTAA 1509
Db   2590 AGCCACCCCTCGTCACAATAA 2610
```

RESULT 15
US-09-184-418C-82
Sequence 82, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 82
; LENGTH: 1496
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=962M751.3; gene=gag
US-09-184-418C-82

Query Match    48.7%;    Score 734.2;   DB 4;   Length 1496;
Best Local Similarity   69.0%;   Pred. No. 8.5e-101;
Best Match
Matches 1020; Conservative 0; Mismatches 458; Indels 1; Gaps 1;

Qy      1 ATGGGTGCGAGAGCGTCGGTATTAAGCGGGGGAGAATTAGATCGATGGGAAAAAATTCGG 60
Db   1850 AGAGTTTTGGCCAGGGGCCAGGGATGAGGCAGTTGCAAGAACTCCGGCAAGAAGGCCAC 1909

Qy     61 CTGCCCCCGGGGAAGAAGAAGCAATACAATGCTGGGGGGGGAGAATTAGATCGATGGGAG 120
Db   1910 GGCAACTTCCGGAACCAGGGAAGATCGTCAAGTGCTTCAAGTGGGGAAAGGAAGGGCAC 1969

Qy    121 CTGGGAGGGGCTTCGCCCTTGGGAGACCCCGAGGGCTGCAAGCAGATC 180
Db   1970 ACCCCCAGGAAACTTGCCGCGGGGCCCCCCCCGGAAGAAGAGGCTGGAAGTGCAAGGAAGGA 2029

Qy    181 ATGGAAGAGCTGCAGCCCCTGCAGCCGCCGCTCGCAGAAGGACTGCGGCCAGC 240
Db   2030 CACCCAAAATGAAAGATTGTACTGAGGACACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCC 2089

Qy    ... (sequence alignment continues) ...

Qy   1489 AGCGGCCCCCTGACCCAGTAA 1509
Db   2252 AACGACCCCTTCGTCACAGTAA 2272

```
Qy    241 ACGGTGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTCCCGGACACCAAGGAGGGC 300
          | || |||||| ||  |    |||||| ||   | |  ||||| ||| || |||||||||
Db    241 ACAGTAGCAAACTCTTATTGTGTACATGAAAAAGATAAAGGTACAGACACCAAGGAAGCC 300

Qy    301 CTGGACAAAGATCGAGGAGGAGGACGAGAACAAGAGTCCCAGCAGAGACACCCAAGGAG 360
          |   ||| || | ||| ||| |  ||| |   ||  ||  ||  |||  || |||||| |
Db    301 CTAGACAAAGATAGAGGAGGAGAACAAAAATCAAAAAAATCAAGAGGCCCAAGGAAGCG 360

Qy    361 GCCGACGGCGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCCAGATGGTG 420
          | || |||| |  |||||||| |||  |   |   ||  |||||| |||||| || |||| |
Db    361 ACTGGCCGGAAAGGTTCAGTCAAAATTATCCTATAGTGTCAGAATCTTCCAAGGGCAAAATGGTA 420

Qy    421 CACCAGGCCCATCAGCCCCTGAACGCCTGGGTGAAGGTTCATCGAGGAGGAGGCC 480
          |||||||| ||||| |  ||| | |  ||||| | |||| ||  | ||| ||| |
Db    421 CACCAGGCTGTATATCACCTAGAACTTTGAAATGCATGGGTAAAGAGTAATAGAGAGAGGGT 480

Qy    481 TTCAGCCCCGGAGGTGAATCCCCATGTTCACCGCCCTGAGCGAGGGGGGCCACCCCCAGGAC 540
          |||  |||| || ||| || |  |  ||| ||  | |||| ||||| || ||||  || |
Db    481 TTCAAACCCAGAGGGTAATACCCATGTTTACGCATTATCAGAAGGCCACCCCCACAAGAT 540

Qy    541 CTGAACACGACGATGTTGAACACCGTGGGGGCCGGCCACCCAGGACGTGCAGATGCTGAAAGGAC 600
          || ||  ||  |  ||| |||| || |||| ||||  |  ||||| || ||| | |||||| |
Db    541 CTAAACACCATCTGTAAATACAGTGGGGGACACTCAAGCACCAAGGTTAAAAAGAT 600

Qy    601 ACCATCAACGACGAGGAGCGCCGCGGAGGTCGGACCGCTGCACCCCGTGCAGGCCGGCCCCGTG 660
          ||||| |  ||| |  |||  ||  || || || ||  |  |  || | |||| |  || | |
Db    601 ACCATCAATGAGGAAGCTCAGAATGGGGGACAATCAAGCAGCATGACAAATGTTAAAAGAT 660

Qy    661 GCCCCCGGCCCAGATGCGCGGCCGACGTCCCAGCAGCGACATCGCCGCACAGACCCTG 720
          | |||  || || || | || ||| |||||||| ||| ||  | || |||||||||
Db    661 GCACCAGGCCCAAATTAAGCAGAACCAAGGGAAGTGACATAGCCACTACTGGTTACCCTT 720

Qy    721 CAGGACAGAGATCGCCTGATGAGCAACCCCGTGCCCGGTGGGGCAGCAGCACATCTACAAG 780
          ||||| ||  ||  |  |  ||| |||||| || |||  |||||  ||||| ||  | |
Db    721 CAGGAACAAATAGCATGGATGACGAACCTATTCCAGGTTGGGGAGCACTCTATAAA 780

Qy    781 GACATTCCGCCCAGGGCGCCCCAAGGAGCCCTTCCGCGACACTAGCTGGACGCTTCTTCAAGACC 840
          | || |  ||| ||| ||||| ||||| |  ||  ||| || || || | |  || ||||||
Db    781 AGATGGATAATTCTGGGGTTAAATAGTAAGGAGTCGACCCCTTAGGAGATTCGGTTCTTTAAAACT 840

Qy    841 GACATCGAGATCGTCCTGGGGCCCCAAGGAGCCCCTTCCGCGACACTAGCTGGACGCCTTCTTCAAGACC 900
          | || |||| | ||  | ||| || |||| |  || | | ||  || || |||   ||  || ||
Db    841 GACATAAAACAAGGACCAAAAGGAACCCTTAGGCATTAATCGGTTCTTTAAAACT 900

Qy    901 CTGCGCGCCGAGCAGGCCACCCACCCCAGGAGACCGTGAAGAACTGACCCTGCTGGGTG 960
          |  |  | |||||| |  |  ||  || ||| | ||  || ||||  || |  |  |||
Db    901 TTAAGAGCTGAACAAGCTACACAAGATGT-AAAAATTGGAATGACAGACACCTTGTTGGTT 959

Qy    961 CAGAAACCGCAACCCGACTGCAAGACCAATCCTCGGCCGCGCTCTCGGCCGCCACCCTG 1020
          ||| |  |  || || |  ||||| |||| || || ||  | |  |   ||  ||||
Db    960 CAAAATCGCAACCCAGAATTGTAAGACCATTTTAAGGGCATTAGACCAAGAGTTTG 1019

Qy    1021 GAGGAGATCAATGACCGCCTGCCACCCGGAGTGGGGCGTGGGCGCGCCAAGCCCGGCTGG 1080
           || || | || || ||  ||  | |  || || ||||| |  || | |||||| |  |||
Db    1020 GAAGAAATGATGACAAGCATGTCGCCACTGGTGGGGAACTCGCCAAAGAACAAGACTTTG 1079

Qy    1081 GCCGAGGCGCCGATGAGCCGGCCATGGGGAACGTCAAGGTGCTTCAACTGCCGAAGTGTGAAG 1140
           | ||||| || |||| |||| ||   |||| ||  |||| ||  | |||| || ||| |||
Db    1080 GCTGAAGCAAATGAACCAACCAGTCAACTGTGGCACAAGAGGCCACATTTAGACCAGGAAT 1139

Qy    1141 GGCCCCCGGCCGCCCCAAGAGGGCTCTGGAAGTGCGGCAAGGGGGGCCACCAGTGAAG 1200
           | |||  |  | |  ||  | | || |||| ||  | || | |  | || | ||||
Db    1140 GGCCCTAAAAGAATTGTTAAATGTTTCAACTGTGGCAGAGGGCCACATTAGACCAGGAAT 1199

Qy    1201 TGCCGAGGCGCCCCCCCAAGAGGGCTCGGAAGTCGGGCAAGGGGGGCCACCAGTGAAG 1260
           |  ||| || | | | | | ||  | |   || |  ||| | | |  ||| | ||||
Db    1200 TGCAGGGGCTCTTCTGGGAAAAAAGGCTGTTTGGAAAGATGTGGAAAAGGAACGAAATGAAA 1259

Qy    1261 GACTGCACCGAGCGCGCCAGGCGCAGGCCAACTTCCTGGGGCAAGATCTGGCCCAGCAAGGGGAGG 1320
           || ||  | ||| |  ||  | ||||| ||||  | ||| | |  |||| ||  |||| ||||||
Db    1260 GACTGTACTGAGAGACAAGGCCTAATTTTTTAGGGAAAATTTGGCCTTCCCACAAGGAAGGGGAGG 1319
```

```
Qy   1321 CCCGGCAACTTCCTGCAGAACCGCAGCGAGCCGCGCCCCCCCCACCGGCCCCCCC 1380
          |  | |||||||  || || ||| | | | |  ||    |   | ||  ||  |  |||||
Db   1320 CCGGGGAACTTCCTTCAGAACAAGACCCACGAGCCAACAGCCAACGAGCCGCCCA 1379

Qy   1381 CCCGCCCAGGACGCTTCCCGCTTCCAGCAGGAGCACACACCCCCCGCCCCCAAGCAGCAGGAGGAC 1440
          ||| | |||||||  |||  | || ||||  |  |  || |||| |   || ||| |||||| ||
Db   1380 CCAGCAGGAGAGCTTCAGGTTCAGGAGGACAACCCCTCGCCCCCGGGGGCAGGACGGAGGACGAAAAGAC 1439

Qy   1441 CGGCGAGCCCCTACCCGGAGCGCCCCTCGACCCGGCCCCGCCCTGACCGCCCCTGCGCGAGC 1479
          |  | ||||  | | ||  || |    | | || | |||| | || |||   || | |||||
Db   1440 AAGGAACCCCTTAACTCGCCCTCAAATCACTCTTGGGCAGC 1478
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run On: June 1, 2004, 10:43:01 ; Search time 3669 Seconds
(without alignments)
12281.829 Million cell updates/sec

Title: US-09-475-704A-4
Perfect score: 1509
Sequence: 1 atgggcgcccgccgcagcat.........gcggcccctgagccagtaa 1509
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 27513289 seqs, 1493109076 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

BST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 83.2 | 5.5 | 933 | 28 | CC391778 | CC391778 PUHIJ70TB |
| c 2 | 81.6 | 5.4 | 1219 | 14 | CD500131 | CD500131 CDA43-C08 |
| 3 | 81 | 5.4 | 749 | 13 | BQ744212 | BQ744212 WHB4113 A |
| 4 | 80.6 | 5.3 | 1674 | 10 | BB636712 | BB636712 rockefell |

(The following are summary entries, column "c" indicates complement)

| | | | | | |
|---|---|---|---|---|---|
| c 5 | 80.2 | 5.3 | 746 | 14 | CK215205 | CK215205 FGAS02715 |
| c 6 | 79.8 | 5.3 | 1201 | 13 | BX356664 | BX356664 BX356664 |
| c 7 | 79.8 | 5.3 | 2299 | 11 | AY106831 | AY106831 Zea mays |
| 8 | 78.4 | 5.2 | 1126 | 14 | CK211108 | CK211108 FGAS02294 |
| 9 | 77.4 | 5.1 | 1064 | 14 | CK163121 | CK163121 FGAS01573 |
| 10 | 77 | 5.1 | 786 | 29 | CG351928 | CG351928 OG3AG53TH |
| c 11 | 76.8 | 5.1 | 1138 | 14 | CK163513 | CK163513 FGAS01614 |
| 12 | 76.4 | 5.1 | 759 | 10 | BF259495 | BF259495 HVSMBf001 |
| 13 | 76.4 | 5.1 | 869 | 14 | CKL59167 | CKL59167 FGAS04056 |
| 14 | 76.4 | 5.1 | 1308 | 11 | AY104577 | AY104577 OG2BH19TV |
| 15 | 75.4 | 5.0 | 883 | 29 | CC274385 | CC274385 WHE2495_B |
| c 16 | 75 | 5.0 | 509 | 12 | BG314500 | BG314500 CDA23-C11 |
| c 17 | 74.4 | 4.9 | 1170 | 14 | CD496623 | CD496623 Zea mays |
| 18 | 74.4 | 4.9 | 2598 | 14 | AY103647 | AY103647 Zea mays |
| 19 | 74.2 | 4.9 | 649 | 14 | CB870888 | CB870888 HC15J04w |
| 20 | 74 | 4.9 | 637 | 12 | BI960118 | BI960118 HVSMEn002 |
| 21 | 73.8 | 4.9 | 562 | 12 | BF484304 | BF484304 WHE2321_P |
| c 22 | 73.2 | 4.9 | 636 | 14 | CD932185 | CD932185 GR45.1117C |
| c 23 | 73.2 | 4.9 | 694 | 12 | BJ261281 | BJ261281 BJ261281 |
| c 24 | 72.6 | 4.8 | 575 | 13 | BQ842001 | BQ842001 WHE2987_G |
| 25 | 72.6 | 4.8 | 601 | 12 | BJ290713 | BJ290713 BJ290713 |
| 26 | 72.6 | 4.8 | 652 | 12 | BJ292215 | BJ292215 BJ292215 |
| c 27 | 72.6 | 4.8 | 683 | 14 | CD905815 | CD905815 G468.102P |
| c 28 | 72.6 | 4.8 | 700 | 12 | BJ295604 | BJ295604 BJ295604 |
| c 29 | 72.6 | 4.8 | 712 | 13 | BJ878344 | BJ878344 WHE2990_E |
| c 30 | 72.6 | 4.8 | 718 | 12 | BQ842240 | BQ842240 BQ842240 |
| c 31 | 72.6 | 4.8 | 753 | 12 | BJ297319 | BJ297319 BJ297319 |
| 32 | 72.6 | 4.8 | 1150 | 14 | BJ298801 | BJ298801 BJ298801 |
| 33 | 72.6 | 4.8 | 1197 | 14 | CK161908 | CK161908 FGAS01448 |
| 34 | 72.2 | 4.8 | 718 | 14 | CK161927 | CK161927 FGAS01450 |
| 35 | 72.2 | 4.8 | 718 | 14 | CD938268 | CD938268 OV.109123 |
| 36 | 72 | 4.8 | 708 | 13 | CD938289 | CD938289 OV.109K04 |
| 37 | 72 | 4.8 | 731 | 13 | BJ752847 | BJ752847 AZ04.102I |
| 38 | 71.8 | 4.8 | 1491 | 14 | CD878344 | CD878344 WHE4119_G |
| 39 | 71.6 | 4.7 | 655 | 12 | AY104431 | AY104431 Zea mays |
| c 40 | 70.8 | 4.7 | 1327 | 10 | BJ300074 | BJ300074 BJ300074 |
| 41 | 70.6 | 4.7 | 671 | 12 | BE636532 | BE636532 rockefell |
| 42 | 70.6 | 4.7 | 738 | 13 | BI718023 | BI718023 1011023D0 |
| 43 | 70.4 | 4.7 | 764 | 13 | BU295509 | BU295509 603608565 |
| 44 | 70.2 | 4.7 | 648 | 10 | BQ804819 | BQ804819 WHB3559_C |
| c 45 | 69.6 | 4.6 | 690 | 14 | CD495520 | CD495520 CDA17-A10 |

ALIGNMENTS

RESULT 1
CC391778
LOCUS     CC391778     933 bp    DNA    linear    GSS 19-MAY-2003
DEFINITION  PUHIJ70TB ZM_0.6_1.0_KB Zea mays genomic clone ZMhBTa45BK20,
          genomic survey sequence.
ACCESSION   CC391778
VERSION     CC391778.1    GI:30871868
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 933)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
          Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
          Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other GSSs: PUHIJ70TD
          Contact: Cathy Whitelaw
          TIGR
          9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 301-838-5843
          Fax: 301-838-0208
          Email: whitelaw@tigr.org
          Seq primer: TR

## Left record

FEATURES
  source
    Class: sheared ends.
    1..933
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone_lib="ZMMBTa458K20"
    /clone="ZMMBTa458K20"
    /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 KB"
    /note="CoT selected genomic DNA library"

ORIGIN

Query Match     5.5%;   Score 83.2;   DB 28;   Length 933;
Best Local Similarity   45.7%;   Pred. No. 1;
Matches 329;   Conservative   0;   Mismatches 388;   Indels   3;   Gaps   1;

```
Qy  189 GCTGCAAGCCCCCCC...GTGACAACACGGTGGC 248
Db   17 GCCGAGGGCGCGG...GTGACAGCGAGCGTGTGGT  76

Qy  249 CACCCTGTACTCGCGTGCA...CCAAGGACGAGGCCCCTGGACAA 308
Db   77 CAGGGTGCACAGCACC...CCCAGAACCGGCTGGT 136

Qy  309 GATCGCAGGAGCAGCAA...CCAAGGAGGCCCGACGG 368
Db  137 GGTGCGTGGAGTTCGCGCG...CTACCCGACCATGGT 196

Qy  369 CAAGGTGAGCCAGAACTA...GTGCTGCACCAGGC 428
Db  197 GCAGCTCAGCCGCCACCT...CGGGACTTCCACGGAGTC 256

Qy  429 CATCAGCGCCCCGCCC...CTTCAGCCC 488
Db  257 GGAGGCCACCGAGGAC...GCCCCACTTCAACTT 316

Qy  489 CGAGGTGGATCCCCATGT...CAGGACTT---GAA 545
Db  317 CTACAAGCGGGCCCGAG...GCGGCTTCGGCGGG 376

Qy  546 CACGATGTTGAACACCGTG...CCTGGACCAT 605
Db  377 CGAGGTCCTCTACTACGG...CTCGGCGGGAGGA 436

Qy  606 CAACGAGGAGGACGCCC...GCGGCCC 665
Db  437 CGTGGAGGCGCTCATCG...CTGTGCTGGACGT 496

Qy  666 CGGCCAGGATGGCGCGA...CTGCAGGA 725
Db  497 GGGCCCTTCAAGGACAC...GAAGCTGTCTCCG 556

Qy  726 GCAGGATCGGCCTGGA...CTTACAAGCGGTG 785
Db  557 GTCCATGGGTCGGACAC...CAGCAGCTGCAT 616

Qy  786 GATCATCCTGGGCGCCC...CATCCTTGGGACAT 845
Db  617 GGAGTTCCTCAGGGCCA...GCATCGGGATCAT 676

Qy  846 CCGCCAGGGCCCCAAGG...TTCTTCAAGACCCTGCC 905
Db  677 CCAGATCGTGCGGCCGCT...GAGATCCTTCCG 736
```

## Right record

AUTHORS      Kingsley,D.M., Peichel,C., Balabahdra,S., Grimwood,J., Dickson,M.,
             Schmutz,J. and Myers,R.M.

TITLE        Expressed sequence tags from Gasterosteus aculeatus

FEATURES
  source
    1..1219
    /organism="Gasterosteus aculeatus"
    /mol_type="mRNA"
    /strain="Salinas river, CA"
    /db_xref="taxon:69293"
    /clone="CDA43-C08"
    /sex="mixed male and female"
    /tissue_type="heads and internal organs combined"
    /dev_stage="adult"
    /clone_lib="SHGC-CDA"
    /note="Vector: lambda ZAP Express/pBK-CMV; Site_1: EcoRI
    (5' adaptor); Site_2: XhoI (3' linker primer); The mixed
    organ cDNA library was generated using the ZAP-cDNA method
    by Stratagene. First strand cDNA synthesis was primed with
    a a 50 bp linker primer containing an oligo dT sequence
    preceded by a synthetic XhoI site. 5 prime adaptors were
    used containing an EcoRI cohesive end. The finished cDNAs
    were inserted in to the ZAP express vector
    unidirectionally in the sense orientation with respect to
    the lacZ promoter of pBK-CMV. An amplified library was
    prepared from approximately 3 million primary clones in
    the lambda ZAP Express vector. In vivo excision was then
    used to generate individual pBK-CMV phagemid clones for
    EST sequencing."

ORIGIN

Query Match     5.4%;   Score 81.6;   DB 14;   Length 1219;
Best Local Similarity   47.7%;   Pred. No. 1.7;
Matches 373;   Conservative   0;   Mismatches 394;   Indels   15;   Gaps   4;

```
Qy  618 CGCCGGAGTGGGGACCGGC...GGCCCCGGCCCAGATGCG 677
Db  917 CGGCTGCGGTGGGCGGACTA...CAGCGCTCGCTGGG 858

Qy  678 CGACCCCGCGGCACGGAC----ATCGCGCGGCGCCA 731
Db  857 CGACTACGGCGGCGGCGACC...GCTACGGCTACGGCGGA 798

Qy  732 CGCCTGGGATGACCAGCCAA...CATCTACAAGCGGTGATCAT 791
Db  797 CGCCCTGCTCACGCCGTCTG...TACGGCTACGGCGGCTC 738

Qy  792 CCTGGGCCTGAACAAGATTCG...CAGCCCCGTGAGCATCCTGGACCA 851
Db  737 GCTGCGGCGACTACGGCGACC...GACTACGGCGGCGGA 678

Qy  852 GGGGC----CCCAAGGAGCCCTT...CGGTTCTTCAAGACCCTGC---G 905
Db  677 CGCCCTGCTCAACGGCTCGCT...GGCGGCCCTGCTCAA 618
```

```
RESULT 3
LOCUS       BQ744212                749 bp    mRNA    linear   EST 17-JUL-2002
DEFINITION  WHE4113_A01_A01ZS Wheat salt-stressed root cDNA library Triticum
            aestivum cDNA clone WHE4113_A01_A01, mRNA sequence.
ACCESSION   BQ744212
VERSION     BQ744212.1  GI:21890999
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   Anderson,O.D., Akhunov,E., Chao,S., Crossman,C., Deal,K.,
            Dvorak,J., Lazo,G.R., Pham,J., Rausch,C.J., Wilson,C. and Woo,J.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Salt-stressed root cDNA library
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oanderson@pw.usda.gov
            Sequences have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: SK primer.
FEATURES             Location/Qualifiers
     source          1..749
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE4113_A01_A01"
                     /clone_lib="Wheat salt-stressed root cDNA library"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid
                     pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Hydroponic
                     plants grown to full tillering stage were treated with 150
                     mM NaCl for either 12 hours or 7 days. Root tissues of the
                     plants subjected to both types of treatment were collected
                     separately at University of California, Davis (E. Akhunov
                     and K. Deal in J. Dvorak's Lab). Total RNA was prepared
                     separately from the two samples (12h and 7day treatments),
                     and equal amount of RNA was then pooled. PolyA RNA was
                     purified from the pooled RNA, a cDNA library was made, and
                     the cDNA clones were in vivo excised to give pBluescript
                     SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak)
                     at the University of California, Davis. Colony plating,
                     plasmid DNA preparations and DNA sequencing were performed
                     in the OD Anderson lab (all other authors)."
                     /tissue_type="Roots"
                     /dev_stage="Full tillering"
                     /lab_host="E. coli SOLR"
ORIGIN
```

```
Query Match    5.4%;  Score 81;  DB 13;  Length 749;
Best Local Similarity   48.4%;
Matches 293;  Conservative  0;  Mismatches 300;  Indels 12;  Gaps  2;

Qy   269 CCCGGCCATCGAGGTCCGCGGACACCCAGAGGAGGGCCTTCAGCCCCCCGAGGTGATCCCC 328
             ||| |||||||| | ||||| ||||| ||||  |||| ||||| ||||||||||||| |||
Db     8 CCTTCAGCCTCGTCGAGACGAGAGATCCCTTGCCAACTTCACCAAGCTTCCTTCGCCG 67

Qy   329 AGTCCCAGCAGAGAGAGACCAGCAGCGACCGAAGGAGGCCGACGGTGAGCCAGAACTACC 388
             |||| || ||| |||| ||||| ||| ||| |||||| |||||| || ||||||| |||
Db    68 GCTTCAACCAGCGCGCCCCAGCTCCCCAGCGCGACGAGGTCGAGGAGGCC 127

Qy   389 CCATCGTGCAGAACCTGCAGGGAGGGCCAGATGTGCACCAGGCCATCAGCCCCCGCACCGTGA 448
             || |  ||| ||||| ||||| |||||||| ||  || |||||| |||| ||||| |||
Db   128 GCGAGGCCGAGAGAGGAGCGGCGCCCCCGGGAGGGCCCACCGGCCGCTCCCACC 187

Qy   449 ACGGCCTGGG-----TGAAGGTGATCGAGGAAGGGGCTTCAGCCCGTGATCCCGTCC 505
             ||| |||  |    |||||||| ||||||| ||||||||| ||||| ||||||||
Db   188 ACCACCAGGGCTTCCGCGCGCAACCAGCAGAACACGCGCCCTTCCGTGCATCCAGC 247

Qy   506 TCACCGCGCCCTGAGCGAGGGGCGCGCCCACCCACCAGAGGACGAACACCGATGTTGAAACACCGTGG 565
             ||||||| ||||| |||| || ||| ||||||||| || || ||||| ||| ||||||||||| ||
Db   248 CCGACTGGAACCATCCTCGAGACAGATCCCCTTGCCAACTTCACCAAGCTTCCTTCGCCG 307

Qy   566 GCGGCCACCAGGCGCCCATGCAGAATGCTGAAGGACGACCATCAACGACGAGGAGGCCCGAGT 625
             ||| ||| ||||||||| ||| ||| |||||| |||| |||||||| || |||||| ||||
Db   308 TCGCCGACCAGCGCCCGCGAGGGACCTGCTTCGTCGCGTGCGCGCCCGGTGCGGGGCCCT 367

Qy   626 GGGACCGCGCTTGCACCCCGGTGCAGGGCGCGCCCCCGGCGCCGGTCGCGCGGACCCCC 685
             |||  ||||||| |||||| |||| || |||||||  |||||| || | || ||||||
Db   368 ACGACCGGCTCACCCCAAGCAGACCGCGCCCGCGCCCCGCGTTCAAAGAACCGCCAGTTCT 427

Qy   686 GCGGCAGCGGCAGCATCGCCGGCGCCGCCACCAGCGACCCTGC---------AGGAGCAGATCGGCGCT 736
             ||| |||| |||| ||| ||||| |||||||||| |||||||         || ||||| | | |||
Db   428 TCAAGATCACCACCACCAGACCCCATCATTCCGCCGCTCGCCGAGGAGGAGCAGGCCA 487

Qy   737 GGATGACCAGCAACCCCCCGTGCCCGTGGGGGACGCATCTACGAGGGTGGATCAGCGGG 796
             || |||| ||||| |||||| ||||| || || ||||| | || ||| || |||||
Db   488 CCGTCTTCGCCACGGACGCCATCCTGCGCGCCCTCATGTGCACGCCGCCAGCAATCCTC 547

Qy   797 GCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCACATCCTGGACATCCGACCAGGCGG 856
             |||||||||||| | || || | ||||| || || |||||| | |||||||| |||||||
Db   548 CCTCGGGACCATTGTCTCGGCAGCGCGCCTCGGCAACAAGCTCTTCTTCGACACGAGGGCT 607

Qy   857 CCAAG 861
             |||||
Db   608 CCCAG 612
```

```
RESULT 4
LOCUS       BE636712                1674 bp    mRNA    linear   EST 03-JAN-2002
```

DEFINITION   rockefeller.0.42 Mastigamoeba balamuthi lambda ZAP II Library
      Mastigamoeba balamuthi cDNA similar to vacuolar ATP synthase
      subunit B (EC 3.6.1.3), mRNA sequence.
ACCESSION   BB636712
VERSION    BB636712.1   GI:9919823
KEYWORDS    EST.
SOURCE     Mastigamoeba balamuthi
  ORGANISM   Mastigamoeba balamuthi
      Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 1674)
  AUTHORS    Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
      Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
      Philippe,H.
  TITLE      The analysis of 100 genes supports the grouping of three highly
      divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
  MEDLINE    21819461
  PUBMED     11830664
COMMENT     Contact: Muller Miklos
      Laboratory of Biochemical Parasitology
      The Rockefeller University
      1230 York Avenue, New York, NY 10021, USA
      Email: mmuller@rockvax.rockefeller.edu
      Insert Length: 1674 Std Error: 0.00
      PolyA=Yes.
FEATURES         Location/Qualifiers
    source        1..1674
           /organism="Mastigamoeba balamuthi"
           /mol_type="mRNA"
           /strain="ATCC 30984"
           /db_xref="taxon:108607"
           /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
           /note="syn: Phreatamoeba balamuthi"

ORIGIN

Query Match   5.3%;   Score 80.6;   DB 10;   Length 1674;
Best Local Similarity   44.1%;   Pred. No. 2.4;
Matches 615;   Conservative 0;   Mismatches 756;   Indels 22;   Gaps 6;

Qy     59 GCCTCGCGCCCCGGGGGGGGGGGCAAGAAGCAACTACACTCTGAAGCACTTGGTGTGGGCG 118
Db     136 GCCCTCTGGGTCGTCCTCGACACATCAAGAAGCCGCAGTTCGCAGACATCGTCAACCTGA 195
Qy     119 AGCTTGGAAGGGGTTCGCCCTGAACCCCGGCCTTCGCTGGAGAGACCGGCGGCTGCAAGACGA 178
Db     196 CGCTGGGCGACGGCACGGTGCCCGCGCCAGGTGCGTCGTGAGGTGGCCCGGAACAAGGCCG 255
Qy     179 TCATGAAGCAGCTGACGCACCCGCCCTGAGACCGGCAGCTGGCCAGCCTGTACA 238
Db     256 TCGTGCAGGGTCTTTGAGGGGCACGTCGGGCCATCGACGGCCAAGAACACCGCACTGGAGTTCA 315
Qy     239 ACACCGTGGCCACCCGTGTACTCGTTGACGAGTTCGCGCAGCAGCCNAGGA-- 296
Db     316 CGGGGCGACATCATNGCGCATCCCGTCAGCGACCATGCTGGGCGTCTTCAACGGCT 375
Qy     297 -GGCCCTGGACAAGATCGAGGAGGAGCAGACAGAGTCCAGCGAGGACCCGCAGCGCCA 355
Db     376 CGGGGCCGCCGGCGCTCGACAAGGGCCCGGCGTCATGGGCGACAGCATCAGG 435
Qy     356 AGGAGGCCGACCGGCAAGGTGAGCCAGAACTACCCATGTGTGCAGAACCTTGCAGGGCCATCA 415
Db     436 GCCAGCGCCATCAAACCCGTACACAGCGTGTAACCCCGACGGAGATGACCGGCATCA 495
Qy     416 TGGTTGCACCGAGGCCATCGACGCCCCGCACCCTGGGTGAAGTGATCGAGGAGA 475
Db     496 GCGGCCATGCGACACGATGAACTCGAGTCGGCTCCAGAAGATCCGCTGTTCAGCGCGCG 555
Qy     476 AGGGCCTTCAGCCCCGAGGTGATCCCGGCCCANGTTCACCGGCGGCCACCCCG 535
Db     556 CCGGTGCTGCCCACACAGAGATTGCCGCGCAGATTCTGCAGGGAGCGCGCCCTCGTGTGCCC 615
Qy     536 AGGAACCTGAACACGATNGTTGAACACCGTNGGCGGCGGCCACCAGGCCGCCATNGCAGGATGCTCGA 595

Db     616 TCCAGAGCAGCAGAAGGTGCAGGACGACGACAACTTCGCCATCGTGTTCGCGG 675
Qy     596 AGGACACCATCAACGAGCAGAAGGCCGGCCGGGACCGGCCTGCACCCCGTTGCAGCGCGCC 655
Db     676 CCATGGGGTGTCAACAGTTGAGAGACTNGCCCAGCAGTTCTTCGCCGCCAGGTTCGAGGAGAACGGCT 735
Qy     656 CCGTGGCCCCCGCCGGCGAGATGCGGCGGACCCCCGCGGCGGCCAGCGCAGCATGCGCGGCGGCCACCAGCA 715
Db     736 CGATGGAGCGCTGNVCAGCGCTGTTCCTCAACCTCGCCAACCCACCCCGGATCGGCGCATCA 795
Qy     716 CCCTGCAGGAGCAGGATCCGCTGTNTGACCAGCAACCCCCCGGTGCCCGTGGGGCCAGACATCT 775
Db     796 TCACGCCGCGTCTG---GCGCTCACGACGGCCAGACCATTGCTCGCACGGCTGCCGAGATGC 852
Qy     776 ACAAGGCCGTNGGANTCATCCTNGGGCCT-----GAACAAGATCGTGCGGGATGAGTACAGCCCCG 829
Db     853 ACGTNGCTNGTTCATTCTCNGGCCATGTGTCGTACCCGGGAGTNCTCGG 912
Qy     830 TGAGCACTTCTGGACANTCCCCNCAGGGCCCCAAGGAGCCCTTCCGGCGACTRAGGTNGGACCNGCT 889
Db     913 CGGGCCGTNGAGGAGGTGCCTTGGGGCGTNCTNGGTTACCCCGGCTACATNGTNACACGGGATNCTCG 972
Qy     890 TCTTCAAGACCTTGCGGCGCCGAGCAGGGCCCCACCCAGGACGGTGAAGAAACTGATNGACCGGAGA 949
Db     973 CGACCATCTACGACGTGCGCGGGCGTG---TCGAGGGGCAGGAGACNGGCTGAATCAGCGGCAGC 1029
Qy     950 CCCTNGCTNGTNGTGCGAGGAGGGAGTGAGGCCGGCNCAACGTNGAACATNGATNGCAGA 1009
Db     1030 CCATCAACTGTNGCTNGCCGTNGCTCAGCCGTNGTNGATNGAAGAGCCCATNGGGCATNGGGGCT 1089
Qy     1010 GCGGCCACCCTGGAGANGGATGATNGACCGGCTGCCAGGGGCCGTNGGGCGCCCCCACAAGG 1069
Db     1090 ACATCACCGAGGGGCCCAGATCTAGCGTNGCGGGGCGTG---TCGAGGGGCAGGAGAACGGGCTGAATCAGCGGCAGC 1149
Qy     1070 C---CCGCGTNGCTGGCCGGAGGCGGANTNGAGCCGCCAGGGCAACACGCGTNGAACATNGATNGCAGA 1126
Db     1150 CCATCAACGTNGGCTNGCCGTNGCTCAGCCGTNGTNGATNGAAGAGCCCATNGGGGCATNGGGGCT 1209
Qy     1127 AGAGCAACTTCAAGGGCCCCGGGCGGCAAGCGCGCAACGGTCAAGTGCTTCAACTNGCGGGCACGGAGGG 1186
Db     1210 CGCCGACGAGCCACTCGGAGGTGCGTNGACGCGGACTNACGGCCATCGGCGCAAGGAGG 1269
Qy     1187 ACATCGCCAAGAACTGCCGGCGCCCCCGCAAGAAGGGTGCCTGGAGATGGCGGCCAAGGAG 1246
Db     1270 ACACGGATGCCATGAAAGGCTGTCGTCGTGCCGGCCCCTTGTCCCCGAGGACAAGCTCG 1329
Qy     1247 GCCACCAGATGAAAGGACTCACCCCAGCGCCAGGGGCAACTTCCTNGGGCAAGATCTGCCCA 1306
Db     1330 CCCTNGGAGTTCCTCGAGNGATGCAGAAGTTCATCAGCAGGGCTACAAACGAGTCGC 1389
Qy     1307 GCCACAAGGGCGCCGCCCCCGGCCGCAACTTCCTNGCAGAGAACCGGCAGCGGGGCCACCACCG 1366
Db     1390 GCGACATCTACCCAGTCGCTGACANTGGCCTNGGG----CCAGTTCCCCAAGG 1445
Qy     1367 TGCCCACCCGGCCCCCCGCGCGAGGAGCTTCGGGTTCGGAGGAGGACCACCCCCCAAGC 1426
Db     1446 AGCTNCCTCACGGCGTNATCACGGGAGGAGGACCCTCAACAAGTNACTNACGGACGGCCAAGN 1505
Qy     1427 AGGAGCCCAAGGA 1439
Db     1506 AGGAGCACCAGTA 1518

RESULT 5

LOCUS      CK215205/c    CK215205    746 bp    mRNA    linear    EST   09-DEC-2003
DEFINITION   FGAS027158 Triticum aestivum FGAS: Library 6 CAP GATE 1 Triticum
      aestivum cDNA, mRNA sequence.
ACCESSION   CK215205
VERSION    CK215205.1   GI:39621309
KEYWORDS    EST.
SOURCE     Triticum aestivum (bread wheat)
  ORGANISM   Triticum aestivum

## RESULT 5

```
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 746)
  AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Budes,P., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
            Links,M.G., McCarthy,B.L., Monroy,P., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE     Functional Genomics of Abiotic Stress In Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas_ests@cs.usask.ca
            This sequence is the direct result of the Base calling software
            Phred (default parameters). It is the raw base calls. To aid in the
            identification of the high quality insert the software Lucy
            (default parameters) has been run on this sequence. Lucy identified
            the region [13,662].
            Plate: L6B016 row: K column: 15.
FEATURES             Location/Qualifiers
     source          1..746
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /clone_lib="Triticum aestivum FGAS: Library 6 CAP GATE 1"
                     /db_xref="taxon:4565"
                     /note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Crown
                     (50%) and leaf (50%) tissues from wheat cultivar Norstar
                     after short exposure times to low temperature in the light
                     and in the dark. 12 mRNA populations were combined before
                     constructing the library. The first 6 populations were
                     days of growth at 20C from wheat cultivar Norstar after
                     short exposure times to low temperature in the light and
                     in the dark. 12 mRNA populations were combined before
                     constructing the library. The first 6 populations: After 7
                     days of growth at 20, wheat plants were transferred to 4C
                     in the light. 1cm crown sections and green leaf tissu were
                     separately harvested after 1, 3, and 6 hours of low
                     temperature exposure. The last 6 populations: After 7 days
                     of growth at 20C, wheat plants were transferred to 4C in
                     the dark. 1cm crown sections and green leaf tissu were
                     separately harvested after 1, 3, and 6 hours of low
                     temperature exposure. First strand synthesis in this
                     library was done in the presence of methylated dCTP
                     thereby protecting from internal cleavage with NotI. In
                     addition, this library used a primer for second strand
                     synthesis that annealed to an artificial sequence (RNA
                     oligo) added before first strand synthesis. Therefore when
                     sequences from EST generated from this library will be
                     masked for vector and adaptor sequences, an additional
                     masking step will have to be included to mask this RNA
                     oligo that is common to all clones (sequence
                     CGACTGGAGCACGAGGACACTGACTGACTGAAGTAGAAA)."
ORIGIN
Query Match  49.7%;  Score 80.2;  DB 14;  Length 746;
Best Local Similarity  49.7%;
Matches  261;  Conservative  0;  Mismatches  258;  Indels  6;  Gaps  2;

Qy   191 TGCAGCCCGCCCTGCAGACCGGACCACCGGAGCTGCGCAGGCCTGTACAAACACCGTGGCCA 250
             ||| ||  |||     || |||||| ||   ||| ||||||   || |||||||||||
Db   670 TGCTGCTCAAGCTCATCGGCCGCCAAGAGACCCATGGGAGGTCCGCGTTCACACCGGTACT 611

Qy   251 CCCTGTACTGCGTGCACGCCGCCATCGAGGTCCGCGACACCAAGGAGCCCTTGACACGCCG 310
             ||||  | |||  |||||||||  ||||| |||||||||||||||||   |||||||||
Db   610 CCCTTCGTCTGCGCCGACACGGCGCATCGAGG... 551

Qy   311 TCGAGGAGGAGCAGGAACAAGTCCCAGCAGACCCAGCAGAGGAAGAGGCCGAGCCGACGGCA 370
             |||||  || | |||||||  | ||| ||||   ||| |||||||||  |||||||||
Db   610 TCGAGGAGGAGCAGGAACAAGTCCCAGCAGACC...CCGAGCCGACGGCA 551
```

## RESULT 6

```
LOCUS       BX356664                1201 bp    mRNA    linear   EST 05-MAY-2003
DEFINITION  BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI015YB03 3-PRIME, mRNA sequence.
ACCESSION   BX356664
VERSION     BX356664.1  GI:30378083
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI015CA02NP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
                     /clone="CS0DI015YB03"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /db_xref="taxon:9606"
ORIGIN
Query Match  9.6%;  Score 79.8;  DB 13;  Length 1201;
Best Local Similarity  9.6%;
Matches  59;  Conservative  303;  Mismatches  254;  Indels  0;  Gaps  0;

Qy   894 CAAGACCCTGCGCGCGAGCGGCGCACCCAGGAGACGTGAAGAACTGGATGACCGAGACCCT 953
            ::|::::::  ::::::  ::::::::::   :|:|||  |  ||::| ||:||:|||:
Db  1201 VAAVSSSAVVSSSSAVVAASSVAASSSSSSSSSSSVSSVAAAAASATVVAAAASASSSSS 1142
```

```
FEATURES
         source          1..2299
                         /organism="zea mays"
                         /mol_type="mRNA"
                         /db_xref="MaizeDB:638374"
                         /db_xref="taxon:4577"
                         /clone_lib="Maize Mapping Project/DuPont Cornsense
                         library"
                         /note="this sequence is part of a project of EST
                         assemblies resulting from the application of public
                         contigs to seed DuPont contigs; this resource was
                         assembled by DuPont as part of a collaboration for the
                         overgo addressing of BACs in conjunction with the Maize
                         Mapping Project"
ORIGIN

Query Match    5.3%;   Score 79.8;  DB 11;   Length 2299;
Best Local Similarity 43.9%;  Pred. No. 3.1;
Matches 495; Conservative  0; Mismatches 617; Indels 15; Gaps  3;
```

```
RESULT 7
AY106831
LOCUS       AY106831               2299 bp    mRNA    linear   HTC 16-OCT-2002
DEFINITION  Zea mays PCO142079 mRNA sequence.
ACCESSION   AY106831
VERSION     AY106831.1  GI:21209909
KEYWORDS    HTC.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 2299)
  AUTHORS   Gai,X., Lal,S., Xing,L., Brendel,V. and Walbot,V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 2299)
  AUTHORS   Coe,E.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT     If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
            maize cDNA sequences is either Virginia Walbot, Stanford or Pat
            Schnable, Iowa State, then clones may be requested from ZmDB:
            www.zmdb.iastate.edu.
```

RESULT 8

```
Qy    950 CCCGCTGCGTGCGCAGAAACGCCAACCCGACTGCAGACCATCCTGCGCGCYCTGCGCCCCG 1009
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   1607 CCGTGCTGGTGCGACCAGCGCCGTGCCAGGCGGCGGAAGGAGGACGCAGCGGTCCG 1666
```

```
Qy   1010 GCGGCCACCCTGGAGGAGGATGANTGACCGGCCTTGCCCAGGGGTGGGGCGGCCACAAGG 1069
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1667 TGTTCTCAAGATCAACAGGTTCGAGGAGGAGCTTGCGGAGGGCCTTGCCCAGGGAGATGG 1726
```

```
Qy   1070 CCCGGCGTGCTGGCCGAGGGCGANTGAGCCAGGGACACATCATGATGCAGAAGA 1129
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1727 AGGCCGCCGGCGGTGGCGCTTCGACGGCGGCGCCGGCGATCGCCAACCAGGATCAAGGAGA 1786
```

```
Qy   1130 GCAACTTCAAGGGCGCCCCGGCGCCCCCGGCGCGCACCGCGGAGAGGGGGCCACA 1189
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1787 GCCGGTCGTA---CCCGCGTACCGCTTCATCGCCAAGGACCTTCGGCGCCGTGTACCTGA 1843
```

```
Qy   1190 TCGCCAAGAACTGCCGCGCGCCCCCCGGCAAGAAGAGGGGCTGCGTGAAGTGC 1236
          |||||  ||||||||||||||||||||||||||||||||||||||||||||
Db   1844 CCGGCGAGAAGCTCAAGTCCCCGCGGCGGAGGAGTGCAACAAGTGTTC 1890
```

LOCUS        CK211108           1126 bp    mRNA    linear    EST 09-DEC-2003
DEFINITION   CK211108/c
             PGAS022942 Triticum aestivum PGAS: Library 6 CAP GATE 1 Triticum
             aestivum cDNA, mRNA sequence.
ACCESSION    CK211108
VERSION      CK211108.1  GI:39617217
KEYWORDS     EST.
SOURCE       Triticum aestivum (bread wheat)
  ORGANISM   Triticum aestivum
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Pooideae; Triticeae; Triticum.
REFERENCE    1  (bases 1 to 1126)
  AUTHORS    Allard,F., Crosby,W.L., Danyluk,J., Budes,F., Frick,M., Gaudet,D.,
             Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
             Links,M.G., McCarthy,B.L., Monroy,A., Muzak,I., Nilson,D.,
             Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE      Functional Genomics of Abiotic Stress in Wheat and Canola Crops
  JOURNAL    Unpublished (2003)
COMMENT      Contact: Wm L Crosby
             Bioinformatics
             University of Saskatchewan, Department of Computer Science
             1C101 Engineering Building, 57 Campus Drive, Saskatoon,
             Saskatchewan, S7N 5A9, Canada
             Tel: 306 966 1769
             Fax: 306 966 2033
             Email: fgas_est@cs.usask.ca
             This sequence is the direct result of the Base calling software
             Phred (default parameters). It is the raw base calls. To aid in the
             identification of the high quality insert the software Lucy
             (default parameters) has been run on this sequence. Lucy identified
             the region [12,750].
             Plate: L6B001 row: A column: 20.
FEATURES             Location/Qualifiers
     source          1..1126
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /clone_lib="Triticum aestivum PGAS: Library 6 CAP GATE 1"
                     /note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Crown
                     (50%) and leaf (50%) tissues from wheat cultivar Norstar
                     after short exposure times to low temperature in the light
                     and in the dark. 12 mRNA populations were combined before
                     constructing the library. The first 6 populations: After 7
                     days of growth at 20Cs from wheat cultivar Norstar after
                     short exposure times to low temperature in the light and
                     in the dark. 12 mRNA populations were combined before
                     constructing the library. The first 6 populations: After 7
                     days of growth at 20, wheat plants were transferred to 4C
                     in the light. 1cm crown sections and green leaf tissu were

                     separately harvested after 1, 3, and 6 hours of low
                     temperature exposure. The last 6 populations: After 7 days
                     of growth at 20C, wheat plants were transferred to 4C in
                     the dark. 1cm crown sections and green leaf tissu were
                     separately harvested after 1, 3, and 6 hours of low
                     temperature exposure. First strand synthesis in this
                     library was done in the presence of methylated dCTP
                     thereby protecting from internal cleavage with NotI. In
                     addition, this library used a primer for second strand
                     synthesis that annealed to an artificial sequence (RNA
                     oligo) added before first strand synthesis. Therefore when
                     sequences from EST generated from this library will be
                     masked for vector and adaptor sequences, an additional
                     masking step will have to be included to mask this RNA
                     oligo that is common to all clones (sequence
                     CGACTGGAGCACGAGGACACTGACATGGACTGAAGGAGTGAGAA)."
ORIGIN
```

Query Match      5.2%;   Score 78.4;   DB 14;   Length 1126;
Best Local Similarity 48.1%;   Pred No. 4;
Matches 285;  Conservative  0;  Mismatches 301;  Indels  6;  Gaps   2;

```
Qy    124 GAGGGCTTCGGCCCTTGAACCCCCGCCTTGCTGGAGACCCGAGGGGCTGCAAGCAGATCATG 183
          ||||||| |||||| ||||||||||||| ||||| ||||| || |||||||||||||
Db    763 GAGATCACCGCCAACCACCCATGAAAACCTGATGACGACCTCGCGCGGCAGGAGCCAGTTC 704
```

```
Qy    184 AAGCAGCTGCAGCCCGCCCCTGCACACCGGCCACCCGAGGAGGCTGCGGCGCTGTACAACACC 243
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    583 GCAATCAACCGGAGAACTACGAGCTGGGGCTGCCGTGCATCGAGAAGGCCGGCGGTGGGG 524
```

```
Qy    244 GTGGCCACCCTGTACTGCGGTGCACGCGCGGCATCAGAGGTCCGGACACCAAGGAGGCCCTG 303
          || ||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db    643 GGCTACTCCCTCCTCCGGCCACCGCGCTCGGCCATCCCGACGAGCGGCACCCATTCTTGGCCATG 584
```

```
Qy    304 GACAAGATCGAGGAGGAGCCAGAGAACAAGTCCCAGCAGAGACCCAGCCAGCGCCAAGGAGGCC 363
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 GCAGCCATCAGGCCCCGCCACCCGCACGCCTGGGGTGAAGGTGATCGGAGGAGGAAGGCCTTC 483
```

```
Qy    364 GACGGCAAGGTAGGCCCAGAACTACCCATCGTGCAGAAACCTGCAGGGCCAGATAGGTGCAC 423
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    523 CACAAGATCGACTTCCGGCGGAGGCCCCGGCGGCCGGCTGCTGGAGCCGGTCCTTGGCTGGAGGAC 464
```

```
Qy    424 CAGGCCCATCAGGCCCCGCCACCCTGAACCGCCTGGGTGAAGGTGATCGGAGGAGGAAGGCCTTC 483
          |||||||||||||||||||||||||||||||||||||||||||
Db    463 GAGGCCAACCACGGCCACCTTCGACTTCGGTCTTCGTGGACCCGACAAGGACAAGACTACCTC 404
```

```
Qy    484 AGCCCCGGAGGTGATCCCCATCGTTCACCGCGCCCTGAGCGGCGGCGGCGCCACCCCCGGAGGACCTG 543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    583 GCAATCAACCGGAGAACTACGAGCTGGGGCTGCCGTGCATCGAGAAGGCCGGCGGTGGGG 524
```

```
Qy    544 AACACGACGATGTTGAACACCGTGGGCGCCCACCAGGCGCCCATGCAGGATGCTGAAGGACACC 603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    346 AACACGCGTTCTGGAACCGGTCCGGTCGTCGTTCCCCGACGAGCGCCCCAAGTAGACATC 287
```

```
Qy    604 ATCAACGAGGAGGCGCCGCAGTGGGACCCCGGTGCGAGGCGGCCCGGCCC---CGTG 660
          || |||||||||||||||||||||||||||||||||||||||||| ||||
Db    286 CGGTCCTTACCGCGACTTCGTCCTTCGACTCAACAAGGCCCTCGCGCCGGCCCAGCCGGTC 227
```

```
Qy    661 GCCCCCGGCGCCAGANTGCGCGGACACCCCGCGCGGCCACAGCGACCATCGCCGGCGCCACCA 712
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    226 GACGATCTCCCAGCTCCCGTGCGCGACCGGCCATCACCCTTCTGCCTCGCTGCGCCCCA 175
```

LOCUS        CK163121           1064 bp    mRNA    linear    EST 05-DEC-2003
DEFINITION   CK163121
             PGAS015739 Triticum aestivum PGAS: Library 4 Gate 8 Triticum
             aestivum cDNA, mRNA sequence.
ACCESSION    CK163121
VERSION      CK163121.1  GI:38993028
KEYWORDS     EST.
SOURCE       Triticum aestivum (bread wheat)
  ORGANISM   Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.

REFERENCE   1 (bases 1 to 1064)
AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Budes,F., Frick,M., Gaudet,D.,
   Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
   Links,M.G., McCarthy,B.L., Monroy,A., Muzak,I., Nilson,D.,
   Penniket,C., Roach,J.L. and Sarhan,F.
TITLE   Functional Genomics of Abiotic Stress In Wheat and Canola Crops
JOURNAL   Unpublished (2003)

COMMENT   Contact: Wm L Crosby
   Bioinformatics
   University of Saskatchewan, Department of Computer Science
   1C101 Engineering Building, 57 Campus Drive, Saskatoon,
   Saskatchewan, S7N 5A9, Canada
   Tel: 306 966 1769
   Fax: 306 966 2033
   Email: fgas_ests@cs.usask.ca
   This sequence is the direct result of the Base calling software
   Phred (default parameters). It is the raw base calls. To aid in the
   identification of the high quality insert the software Lucy identified
   (default parameters) has been run on this sequence. Lucy identified
   the region [1,834].
   Plate: L4B007 row: C column: 11.

FEATURES      Location/Qualifiers
   source   1..1064
     /organism="Triticum aestivum"
     /mol_type="mRNA"
     /db_xref="taxon:4565"
     /clone_lib="Triticum aestivum FGAS: Library 4 Gate 8"
     /note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Aerial
     parts (crown and leaf) of wheat cultivar Norstar during
     dehydration stress. 8 mRNA populations were combined
     before constructing the library. The first four come from
     removing plants from vermiculite (7 day old plants) and
     incubating them at 20C on the bench without water for 1,
     2, 3 and 4 days. The last four come from plants grown in
     soil in a growth chamber after watering is terminated.
     Four samplings were taken in a two week period; the first
     after wilting was observed and the last, two weeks later,
     consisted of live crown and stem tissue (leaf tissue was
     yellow and dead). First strand synthesis in this library
     was done in the presence of methylated dCTP thereby
     protecting from internal cleavage with NotI."

ORIGIN

Query Match   5.1%; Score 77.4; DB 14; Length 1064;
Best Local Similarity   48.0%;
Matches 284; Conservative 0; Mismatches 302; Indels 6; Gaps 2;

```
Qy    124 GAGGGCTTCGCCCTGAACCCCGGCCTCGCTCTGGAGAACCCCGGAGGGCTGCAAGCAGATCATG 183
              |||  | |||||||       |||||||||||||||||  ||| |
Db    282 GAGATCACCGCCCAACCACCCATGAGAACCTGATGACGACGTCGGCCGGGAGGGCCAGTTC 341

Qy    184 AAGCAGCGATGCTGCAGCCCCCCCCGAGACCGGCACCGAGAGCTCGCAGCCTGTACACACC 243
                  |||  |||||| ||||||||||||| |  ||||||  |||||  |||||||| ||||
Db    342 CTCAACATGCTGCTCAAGCTCATCCGGCGCCCAAGAGAGACCCATGGAGATGGGGTCTACACC 401

Qy    244 GTGGCCACCCTGTACTCGGTGCACGCGCCGGCCATCGAGGTCCGGCGACCAAGGAGGCCCTG 303
              |||||||||||  ||| || || ||| ||||  |||||| ||||  ||| ||
Db    402 GGCTACTCCCTCCCCCGCCCACCGGCGGCGCCATCCCGACGGCCACCATCTTGGCCATG 461

Qy    304 GACAAGATCGAGGAGGAGCAGAACACAGTCCCAGCCAGCAGGCCGGTCAAGGAGGCC 363
              |||| |||| ||||| ||  ||| |||  || |  ||||| | |  ||| |||||||
Db    462 GACATCAACCGCGAGAACTACCAGCGGGACCGTCCATCGAGAAGGCCCGGCCCGTGGCCG 521

Qy    364 GACGGCCAAGGTGAGCCAGAACTACCACCGGCCGCCCAGATACCCAGATGGTGTGCAC 423
              |||||||||| ||||| ||||  |  | | |||||| |||||||  |||| |||||
Db    522 CACAAGATCGACTTCCGCGAGGGCCCGGCGCTCGACGCGCCTCGTGGAGGAC 581

Qy    424 CAGGGCCATCAGCCCGCCACCTGACACCCTGGGTGAAGAGGTGATCGAGGAGAAGGCCTTC 483
              |||||||  |   || ||  | |||||| |  |||    |||||| || |  ||||  |||
Db    582 GAGGCCAACCACGGCACCTTCGACTTCGTCTTGACCCCTGTAGACGCCCGACCAAGGACAACTACACTC 641

Qy    484 AGCCCGGAGGTGATCCCATAGTTCACCCGCCCTTGAGCGGAGGAGGGCCACCCCCCAGGACCTG 543
              |||||||||||||||||||||| |||||| |||||||| || | || |||||||||||||
Db    642 AACTACCACGACGCGGCCCTCATTGAAGCTTCGTCAAGCTCGGCCTACGAC---  698

Qy    544 AACACGGATGTTGAACACCCTGGGCGGCGGGCCACCAGGCGCCGCCATGCAGATGGTGAAGGACACC 603
              |||||||| ||||||   ||| |||| | || ||||| | ||| ||||||||||||||||
Db    699 AACAGGCTCTGGAAGCCGTCCGGCGTCGTCCCCGACGGCCGCCATGGCCAAGTACATC 758

Qy    604 ATCAACGAGGAGGAGGCCGGCGCCCAGTGGGGACCCCGTGCCACCCGTGCAGGCCGGCCC---CGTG 660
              ||||||||||| |||  |||||  |||   |||| || || ||||  || |||| |||  ||||
Db    759 CGCTACTACCGCGACTTCGTCCTCGACCTCAACAACAAGCGCCTCGCCGGCGTCGCCAGCGCCGGTC 818

Qy    661 GCCCCCGGCCCAGATGCGGCGACCCCCGCCCGACGCCAGCCAGACAGCAGCATCGCCGCGCGGCGCCACCA 712
              || ||| |||||| |||| ||  |||||| ||||  ||| ||  |  |||||||||||||| ||| |||
Db    819 GAGATCTGCCAGCTCCCCTGCGCACGGCATCACCCTCTTGCCGCGCGGCGCCA 870
```

RESULT 10

CG351928/c

LOCUS   CG351928    786 bp    DNA    linear    GSS 26-AUG-2003
DEFINITION   OG3AG53TH ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMa0754J09,
   genomic survey sequence.
ACCESSION   CG351928
VERSION   CG351928.1   GI:34269194
KEYWORDS   GSS.
SOURCE   Zea mays
   ORGANISM   Zea mays
   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
   clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 786)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
   Resnick,J., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohling,T.,
   Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE   Consortium for Maize Genomics
JOURNAL   Unpublished (2002)

COMMENT   Contact: Cathy Whitelaw
   TIGR
   9712 Medical Center Drive, Rockville, MD 20850, USA
   Tel: 301-838-5843
   Fax: 301-838-0208
   Email: whitelaw@tigr.org
   Other GSSs: OG3AG53TV
   Seq primer: TR
   Class: sheared ends.

FEATURES      Location/Qualifiers
   source   1..786
     /organism="Zea mays"
     /mol_type="genomic DNA"
     /strain="B73"
     /db_xref="taxon:4577"
     /clone_lib="ZMMBMa0754J09"
     /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
     methylation filtered genomic DNA library"

ORIGIN

Query Match   5.1%; Score 77; DB 29; Length 786;
Best Local Similarity   46.1%;
Matches 334; Conservative 0; Mismatches 385; Indels 6; Gaps 2;

```
Qy    579 CGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCGGGCGCCCGAGTGGGACGCCGCA 638
              |||||| |||| |||||| ||| || |  || || || ||  | | ||||||| ||
Db    776 CGCCACGCGCCTCTGGAGGCGCAGCGGCCGCCATCCGACGCTGCAGCACCATCG 717

Qy    639 CCCCGTGCAGGGCCGCCGGGCCCCCGTGGCCCCCGCGGCCAGATGCGGGACACGATGCGGACAT 698
              |||||| |||||||  || ||   |||||  ||| | | |||||| ||||| | |||||||
Db    716 GGAGGAGGGGCCGAGGCCGCAGCTCGGCCCAGGCGCCGACCGTGCAGGACGGGTGCT 657

Qy    699 CGCCGCGGCGCCACCAGCAACCCTGCGCGGAGGACAGATCGCCTGGATGACCAGCAACCCCCCGT 758
              ||| ||  | ||  || |   ||| | |||||||| ||| | |||| || | |||||||||
Db    699 CGCCGCTCGACGCGCCCTACCGGTGCCGTCCGGGGGTGCTGCTGGAGAAGTTCCCAA 597

Qy    579 CGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCGGGCGCCCGAGTGGGACGCCGCA 638
```

Query Match   5.1%; Score 77; DB 29; Length 786;
Best Local Similarity   46.1%;
Matches 334; Conservative 0; Mismatches 385; Indels 6; Gaps 2;

RESULT 11
CK163513
LOCUS       CK163513               1138 bp    mRNA    linear   EST 05-DEC-2003
DEFINITION  CK163513 FGAS016142 Triticum aestivum FGAS: Library 4 Gate 8 Triticum
            aestivum cDNA, mRNA sequence.
ACCESSION   CK163513
VERSION     CK163513.1  GI:38993817
KEYWORDS    EST.
SOURCE      Triticum aestivum
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 1138)
AUTHORS     Allard,F., Crosby,W.L., Danyluk,J., Budes,P., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
            Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE     Functional Genomics of Abiotic Stress In Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Wm L Crosby
            Functional Genomics
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas_est@cs.usask.ca
            This sequence is the direct result of the Base calling software
            Phred (default parameters). It is the raw base calls.To aid in the
            identification of the high quality insert the software Lucy

(default parameters) has been run on this sequence. Lucy identified
the region (8, 827).
Plate: L4B008  row: F  column: 04.

FEATURES             Location/Qualifiers
     source          1..1138
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone_lib="Triticum aestivum FGAS: Library 4 Gate 8"
                     /note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Aerial
                     parts (crown and leaf) of wheat cultivar Norstar during
                     dehydration stress. 8 mRNA populations were combined
                     before constructing the library. The first four come from
                     removing plants from vermiculite (7 day old plants) and
                     incubating them at 20C on the bench without water for 1,
                     2, 3 and 4 days. The last four come from plants grown in
                     soil in a growth chamber after watering is terminated.
                     Four samplings were taken in a two week period; the first
                     after wilting was observed and the last, two weeks later,
                     consisted of live crown and stem tissue (leaf tissue was
                     yellow and dead). First strand synthesis in this library
                     was done in the presence of methylated dCTP thereby
                     protecting from internal cleavage with NotI."
ORIGIN

Query Match     5.1%;  Score 76.8;  DB 14;  Length 1138;
Best Local Similarity   47.9%;  Pred. No. 6.3;
Matches  283;  Conservative  0;  Mismatches  303;  Indels  5;  Gaps  2;

Qy    759  GCCCGTGGCGACGACATCTACAAGACGTGGATCATCCTGGGCCTGAACACAAGATCGTGCGGAT  818
Db    596  GCCGGTGAGCCGTGGGTGCCGCGCGGTGGTCCGCCGCGCCCCCGCCAGCAGCCAAGCGGC    537

Qy    819  GTACGAGCCCGTGAGCATCCTGGACATCCGCCGCGGCGGCCCCAAGGAGACCCTTCCGCGGACTA  878
Db    536  CAGCAGCTTCCCGTTCCAGAGCTTCGGGGCGCCGCGGCGGCGAATGGGTGGGGGCCAGGAAGT  477

Qy    879  CGTGGACGCCGTTCTTCAAGACCTGCCGCGCAGGCGGCCACCAGGACGTGAAGAACTG  938
Db    476  GGAGGAGGAGAATGCCGGCGGCCTGCTCGCCGGCGTGCTGCGGGGCCA---AGGACAGCAGGAGTA  420

Qy    939  GATGAACCGAGACCCTGCTGGTGTGCAAGAACAGCCAACCCGGACTGCAAGACATCCTCGCGCGC  998
Db    419  CCTGAGAGGTGGGGGAAGCTCGTGGCGGCGTCAACCCGC---GCCTGCCGGTGGCGCGGGCC    363

Qy    999  TCTCGGGCCCCGGGCGCCACCCTGGAGCAGGAGGATGAGGACCGGGGCGTGGGGCCCC  1058
Db    362  CGGGCTGGACGTGGCCGGGCGCCGCCGCCTTCATCGGGCCGCACGCGCCGTGGGC    303

Qy   1059  CGGCCCAAGGCCGGGCGTCGTGGGCCGGAGGCCGGATGAGCCAACACCGTGAACATCAT  1118
Db    302  GTCGGCGGCGGCGCCGGTCGACGTGGCGGGCGCGGCTGCGCGGCGGTGAACACG    243

Qy   1119  GATGCAGAAGCAGCAACTTCAAGGGCCCCCCGGGCGGCAAGCGTGCTTCTTCAACTGCGGCAA  1178
Db    242  CGGCGCGCTGGGGCATGCTGTTCGAGCGTGTGCGGCAACTGCGGCCGGCCTTTACCGCAAGCT  183

Qy   1179  GGAGGGCCACATCCGCCAAGAGAACTGCGCGGCCGCGCCCCGGCTTGTGGAAGTGCGCG  1238
Db    182  GCAACGGAGCATCGAGGCGGCGCTGGAGCGCCCGAGACCGGTGGGAGAGGAGCGGGGA  123

Qy   1239  CAAGGAGGGCCACCAGATGAAAGGACTGCACCGAGGGCCAGGCCAACTTCCTGGGCGAAGAT  1298
Db    122  CGTGTTCCGGACCAAGGTGGCCTTGCTCCTGCTGCCGCTCCGGCGGTCCGCCCAGTT    63

Qy   1299  CTCGC  1303
Db     62  CAGGC   58

Qy    124  GAGGGCTTCGCCCTGAACCCCGGCCGCTCGGAGAACCCCGGAGGGCTGCAAGCAGANTCATG  183
Db    232  GAGATCACCGCCAACCCACCATGGAACCTGATGAGGACGAGGTCGGCGGAGGACCAGTTC  291

Qy    184  AAGCAGGTGCACGCCCCGCCCTGCAGACCGGCCACCGGAGGAGTCGGCCAGCCTGTACAACACC  243
Db    292  CTCAACATGCNCGCTCAAGCTCATCGGGCGGCCAAGAAGACCATGAGAGATCGGCGTCTACACC  351

Qy    244  GTGGCCACCCTGTTACTTCGGTGCACGCCCCGGCCATGCGAGGTTCGCGCACACAACAAGGCC  303
Db    352  GGCTACTNCCCTGCCCACCCCGCCATCCCCACACGACGGCCACCATTCTTGGCCATG  411

Qy    304  GACACAGGATCGGAGGAGGAGGCAGCGAACAAGTGCCCAGCACGAAGACCCCGGCCGCCAAGGAGCC  363
Db    412  GACATCAACCACGGACGGCCTCATGAAGCTCGTCAAGGTCGCCGGGCGCTCCTGGGCTAGCAC--  471

Qy    364  GACGGCAAGGTGAGCCAGAACTACCCCATCGTGGGCTCCGTGCATCGAGAAGGTGCAC  423
Db    472  CACAAGGATCGACTTCCCGCGAGGCGCCCGGCGCTCCGTGTCCCGTCTCCTGGACAC  531

Qy    424  CAGGCCCATCAGCCCCCGCACCCTGAACGCCCTGGGTGAAGGTGATCGAGGAGGAAGGCCTTC  483
Db    532  GAGGCCAACCACGGGACCTTCGAACTYCGTTCGTTGGACGCGGACCAAGGACAACTACCTC  591

Qy    484  AGCCCCGGAGGTGATCCCCATGTTCACGGCCCTTGAGCGAGGAGGCGGCCACCCCCCAGGGACCTG  543
Db    592  AACTACCACGAGCGGCTCTCATGAAGGCTCGTCAAGGTCGCCGGCGCACCGCTGCGAC--  648

Qy    544  AACACGGATGTTGAAACACCGTGGGCGGCCGGCCCACCAGGCCCGCCATGCAGAGATGCTGAAGGAGCACC  603
Db    649  AACACCCTCTGGAAACGGCTCCCCGTGCTCCCCGGAAAGCGCCCCCATGCGGCAAGTACATC  708

Qy    604  ATCAACGNAGGAGGCCCGCGCGAGTGGGACCCGCTGCACCCCCGTGCAGGCCGGC--CCGTCG  661
Db    709  CGCTACTACCGGCGACTTCGTCCTCGAGCTTCAACAGGCCCTCGCCCGNCGACCAGCGCGTCG  768

Qy    662  CCCCCGGCCGCGATNCGCGGACCCCCGGCGCGCCGCGACATCGCCGCGGCCCACCA  712
Db    769  AGATCTGCCAGTCTCCCCGTCGGCGACCGGCATCACCCTCTGCGCCGCGCGCCA  819

RESULT 12
BF259495
LOCUS       BF259495                759 bp    mRNA    linear   EST 22-OCT-2001

```
DEFINITION  HVSMBf0019H19f Hordeum vulgare seedling root EST library HVcDNA0007
            (Biolated and unstressed) Hordeum vulgare subsp. vulgare cDNA
            clone HVSMBf0019H19f, mRNA sequence.
ACCESSION   BF259495
VERSION     BF259495.2  GI:13120022
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
ORGANISM    Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 759)
AUTHORS     Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D.,
            Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W.,
            Fenton,R.D., Oates,R. and Main,D.
TITLE       Development of a genetically and physically anchored EST resource
            for barley genomics: Morex unstressed seedling root cDNA library
JOURNAL     Unpublished (2001)
COMMENT     On Nov. 16, 2000 this sequence version replaced gi:11188608.
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence stop: 758.
            Total hq bases = 388
            Email: rwing@clemson.edu
            Fax: 864 656 4293
            Tel: 864 656 7288
            100 Jordan Hall, Clemson, SC 29634, USA
            Clemson University
            Clemson University Genomics Institute
            Contact: Wing RA
FEATURES    Location/Qualifiers
    source  1..759
            /organism="Hordeum vulgare subsp. vulgare"
            /mol_type="mRNA"
            /cultivar="Morex"
            /sub_species="vulgare"
            /db_xref="taxon:112509"
            /clone="HVSMBf0019H19f"
            /tissue_type="Seedling root"
            /lab_host="TJC121"
            /clone_lib="Hordeum vulgare seedling root EST library
            HVCDNA0007 (Biolated and unstressed)"
            /note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI.
            Seeds were surface sterilized then germinated under axenic
            conditions in the dark at room temperature on filter paper
            with water, nystatin and cefotaxime in covered
            crystallization dishes. Five-day old seedling roots were
            then harvested, total RNA was prepared, poly(A) RNA was
            purified, one primary unamplified cDNA library was made,
            and 1 million pfu were in vivo excised to give pBluescript
            SK(-) cDNA phagemids. These steps were performed in the TJ
            Close laboratory at the University of California,
            Riverside (Choi, Close, Fenton). Phagemids were plated and
            picked at the Clemson University Genomics Institute (CUGI)
            (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
            preparations, DNA sequencing and sequence analysis were
            performed at CUGI (Wing, Yu, Frisch, Henry, Simmons,
            Oates, Rambo, Main). The sequence has been trimmed to
            remove vector sequence and contains a minimum of 100 bases
            of phred value 20 or above. For more details on library
            preparation and sequence analysis see
            http://www.genome.clemson.edu/projects/barley/. To order
            this clone see http://www.genome.clemson.edu/orders Also
            see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
            Genetically and physically anchored EST resources for
            barley genomics. Barley Genetics Newsletter 31:29-30.
            (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
ORIGIN
```

```
Qy   Query Match   5.1%;  Score 76.4;  DB 10;  Length 759;
     Best Local Similarity  47.8%;  Pred. No. 6.5;
     Matches 283; Conservative  0; Mismatches 303; Indels  6;  Gaps  2;

Qy   124 GAGGGCGTTCGCCCTGACCCCGCGCCTGCTGGAGACCGCGGGGCTCAAGCAGATCATG 183
```

```
Db    13 GAGATCACCGCCAACCACCCATGGAACCTGATGACGACGTCGGCGGACGCGAGGGCCAGTTC 72
             ||||    |||||  || |  ||| | |||| | |   | | |||| |||||  | ||
Qy    13 GAGATCACCGCCAACCACCCATGGAACCTGATGACGACGTCGGCGGACGCGAGGGCCAGTTC 72
```

```
Db   184 AAGCAGCTGCAGCCCGGCCCTTGCAGACCGGCCGCCGGAGGAGCTGCGCAGGAGGCC 243
Qy   133 GGCTACTCCCTGCTCGGCCACCGCGTCCCCGACCGCACGCCACCCGCAGTTTGGCCATG 192
Db    73 CTCAACATGCTGCTCAAGCTCATCGGCGCCGCCAAGAAGAGACCATGGAGATCGGCGGCTTCACACG 132
Qy   244 GTCGCGCACCCTGTACTGCGGTGCACGCCGGCGATCGAGGTTCCGGCGACCAAGGAGGCCCTG 303
Db   304 GACAAGGATCGAGGAGGACGAGCAGAACAAGTCCCAGCAGAAGAGACCCAGCAGGCCAAGGAGGCC 363
Qy   193 GACATTCAACGCGATCAGGGAACTACGAGCGTGGGGCCTGCGATCGAGATCGAGCAGTTGTGCAC 252
Db   364 GACGCGCCAAGGTGCAGCGCCAGAGAACTACCCCCCATCGTGTGCAGAACCTGCACAGATCGTGCAC 423
Qy   253 CACAAGGATCGACTTCCGCGGAGGGCCCCGGCGGTCCCCGGTCCTGGACGCCCTCCGGAGGAC 312
Db   253 CAGGGCCATCAGCCCCGGCACCCTGAACGCCTTGGGGTGAAGGTGATCGAGGGAAGGCCTTC 483
Qy   424 CAGGCCATCAGCCCCCGGCACCCTGAACGCCTTGGGGTGAAGGTGATCGAGGGAAGGCCTTC 483
Db   313 GAGGGCCAACCACGGGACCCTCGACTTCGTCTTCGTGTGGACGCGCGACAAGGACAACTACCTC 372
Qy   484 AGCCCCGGAGGGTGATCCCCATGTTCACCGCCCTGAGCGGAGGGGCCCACCCCCCCAGAGACCTG 543
Db   373 AACTACCACGGAGCGCCCTCATGAAGCTCGTCAACGGTGGCGGGGGGGGCTTCCTCGGCTACGAC--- 429
Qy   544 AACACGGATGTTGAACACCGTGGGGCCGCCACCAGGCGCGCCAGATGCTGAAGGAGGACACC 603
Db   430 AACAACCCTCTGGAACGGCTCCGGCTCGCGGCCCCCCCCATGCCAAGGTACATC 489
Qy   604 ATCAACGAGCGAGGGCCGCCGGAGTGGGACCGGGCCTGCAGCCCCGTGCAGGCCGGGCCC---CGTG 660
Db   490 CGCTACTACCCGCGACTTCGTNCTCGACCAACAAGGCCCTCGCCGACCAGGCGGCGTN 549
Qy   661 GCCCCCGGCCCAGGATGCGGGACCCCCCCGGGGGGGGCGGGACCATCCGCCGGGGGCCACCA 712
Db   550 GAGATCTGCCCAGCTCCCCGTTCGGGCAGGGCCATCACCCTCTCCCGCGCGCGGCA 601
```

```
LOCUS       CK159167              869 bp    mRNA    linear   EST 05-DEC-2003
DEFINITION  PGAS040564 Triticum aestivum FGAS: TaLt5 Triticum aestivum cDNA,
            mRNA sequence.
ACCESSION   CK159167
VERSION     CK159167.1  GI:38985053
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
ORGANISM    Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 869)
AUTHORS     Allard,F., Crosby,W.L., Daryluk,J., Budes,P., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
            Links,M.G., McCarthy,B.L., Monroy,A., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.L. and Sarban,F.
TITLE       Functional Genomics of Abiotic Stress in Wheat and Canola Crops
JOURNAL     Unpublished (2003)
COMMENT     Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas_est@cs.usask.ca
            This sequence is the direct result of the Base calling software
            Phred (default parameters). It is the raw base calls. To aid in the
            identification of the high quality insert the software Lucy
            (default parameters) has been run on this sequence. Lucy identified
            the region [128,636].
```

RESULT 14

LOCUS          AY104577          1308 bp     mRNA     linear     HTC 16-OCT-2002
DEFINITION     Zea mays PCO137756 mRNA sequence.
ACCESSION      AY104577
VERSION        AY104577.1  GI:21207655

FEATURES       Location/Qualifiers
source         1..869
               /organism="Triticum aestivum"
               /mol_type="mRNA"
               /cultivar="Wheat line PI 178383"
               /db_xref="taxon:4565"
               /clone="DH5 alpha"
               /lab_host="Triticum aestivum FGAS; TaLt5"
               /clone_lib="Triticum aestivum FGAS; TaLt5" (suppression
               subtractive hybridization) cDNA library from Genotype
               /note="Organ; Crown; Vector: pGEM-T; SSH (suppression
               PI178383 cold hardened at 2 C for 21 days and 49 days
               (equal amount of cDNA pooled together before subtraction,
               tester) and subtracted against genotype Norstar cold
               hardened at 2 C for 1 day (24 H)(driver). Modified Smart
               cDNA (Clontech)priming and non-directional cloning"

ORIGIN
Query Match    5.1%;  Score 76.4;  DB 14;  Length 869;
Best Local Similarity   45.2%;  Pred. No. 6.7;
Matches 281;  Conservative   0;  Mismatches 341;  Indels   0;  Gaps   0;

Qy   663  CCCCCGCCCAGATNCGGGGACCCCCCGGCGCAGCGAACCCCGTGGGCGGCACCGCCCTTGCA 722
               ||||||||| |||||||| |  || ||||| | |  || ||||| | |||||
Db   780  CCACAACCACCAACGACGACAACAGCCACACATCGCCGCCGCCAGCACCCTTGCA 721

Qy   723  GGAGCAGGATCGCCTGGGATGGACCAGCAACACCCCCGTGGGCGGACCATCTACAACGG 782
               ||||||  |||||||||||||| || |||||| ||||| ||| |||||||||
Db   780  CCACAACCACCAACGACGACAACAGCCACACATCGCCGCCGCCAGCACCCTTGCA 721

Qy   783  GTGGGATCATCTCTGGGCCTGACAAAGATCGTGCGGGGATTCACGCCCCGTGAGCATTCTGGA 842
               |||||||||| | ||||| |||||||||||||||||||||||||||||||||||
Db   660  CAACCAACCACCAACGACGACAACAGCCACACACAACAACAACAACAACAACAACAACA 601

Qy   843  CATCCGCCAGGGCCCCTTCCGCGGACTACGTGCACCGGCTTCTTCAAGACCCT 902
               || ||||||||||| |||||||||||| |||||||||||||||||
Db   600  CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 541

Qy   903  GCGCGCCGAGCGAGGCCACCCAGCAACGTGAAGAACTGGGATGACCTGCTTCGTGGTNCA 962
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   540  CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAA 481

Qy   963  GAACCCCGACTGCAAGACCATCCTGCGCGGTCTTCGGCCCCGGCGCCACCCTTGCA 1022
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   480  CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACGACGA 421

Qy   1023 GGAGATGATGACCGGCCCTGCCAGGGCGGCGGGGGGCCCCGGCCCGCCGGTGCTGGC 1082
               |||||||||||||||||||||||||||||||||||||||||||||||
Db   420  CGACGACGACGACGACGACAACAGCGACAACGACGACAACAACAACAACAACACAA 361

Qy   1083 CGAGGCGCATGAGCGCCAGGCCCAACAGCGTGAACATCATGATGGAGAAGAGCAACTTCAAGCG 1142
               |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   360  CAACCAACAACAACGACGACAACAGCCACACACAACAACAACAACAACAACGA 301

Qy   1143 CCCCCGGCGCAAGTCCTTCAACTGCGGCCACCATCGCCGCAAGGAGAACTG 1202
               ||||||||||||||||||||||||||||||||||||||||||
Db   300  CAACAACCACCAACGACGACAACAGCCACACACAACAACAACA 241

Qy   1203 CCGCCGCCCCCCGGCCAAGAGGCTGTGGAAGTGCGGCGGCCACCCAGATGAAGGA 1262
               || ||||||||||||||||||||||||||||||||||||||||||||
Db   240  CAACAACCACCAACGACGACAACAACAACAACAACGACAACAACAACGA 181

Qy   1263 CTGCAACCACGACGCGCCAGGCCAAC 1284
               || |||||||||||||||||
Db   180  CAACAACGACAACAACGGCAAC 159

KEYWORDS  HTC.
SOURCE    Zea mays
ORGANISM  Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  1  (bases 1 to 1308)
AUTHORS    Rainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
           Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE      Maize Mapping Project/DuPont Consensus Sequences for Design of
           Overgo Probes
JOURNAL    Unpublished (2002)

REFERENCE  2  (bases 1 to 1308)
AUTHORS    Coe,E.H.
TITLE      Direct Submission
JOURNAL    Submitted (25-APR-2002) Maize Mapping Project, University of
           Missouri, Columbia, MO 65211, USA

COMMENT    If you are interested in getting corresponding physical clones,
           these are publicly available from ZmDB and may be found by BLAST
           searching at MSI, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
           www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
           maize cDNA sequences is either Virginia Walbot, Stanford or Pat
           Schnable, Iowa State, then clones may be requested from ZmDB:
           www.zmdb.iastate.edu.

FEATURES   Location/Qualifiers
source     1..1308
           /organism="Zea mays"
           /mol_type="mRNA"
           /clone_lib="Maize Mapping Project/DuPont Cornsensus
           Library"
           /db_xref="taxon:4577"
           /note="this sequence is part of a project of EST
           assemblies resulting from the application of public
           contigs to seed DuPont contigs; this resource was
           assembled by DuPont as part of a collaboration for the
           overgo addressing of BACs in conjunction with the Maize
           Mapping Project"

ORIGIN
Query Match    5.1%;  Score 76.4;  DB 11;  Length 1308;
Best Local Similarity   44.7%;  Pred. No. 7.2;
Matches 519;  Conservative   0;  Mismatches 601;  Indels   42;  Gaps   4;

Qy   294  GGAGGCCCTGGACAAGATCGAGGAGCGAGCAAGAGCAGAGAAGAGACCCAGCAGGC 353
               ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2    GGACGGCGTGGGAGGAGACGAGATGGGGCGGCCTGAACATCGACGAGCTGATCAGAAGGCGA 61

Qy   354  CAAGGAGGGCCGACGGCCAAGGTGAGCCAGAACTACCCCATGTCGCAGAACTTGCAGGGCCA 413
               ||||||||||||||||||||||||||||||||||||||||||||||
Db   62   CGGGTTCGCGCGGGGCGGTTGTTCCCGGAGCACAAGATACGAGATCGTGAAGAGGCGGCCG 121

Qy   414  GATCGTGCACCAGCGCCATCAGCCCCCGCCCACCCTGAACGGCCTGGGTGGGTGAAGGTGATGGAGGA 473
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122  GAAGCACATCTGCGCATCACCGGGACGCGTGAAACGACGGCGCCGGCGCTGAAGAAGGC 181

Qy   474  GAAGGCCTTCAGCCCCGGAGCGTGATCCCCATGTTCACCGCCCTGAGCGACGCCCACCCG 533
               |||||||||||||||||||||||||||||||||||||||||||||||
Db   182  GGACATGCGCATCCGCGGTGACGACGCGGAGGCCGTCGGGACATCGT 241

Qy   534  CCAGGACCTGAACACGATGTTGAACACGTGGGCGGCCACCAGCCGCCCATGCAGATGCT 593
               |||||||||||||||||||||||||||||||||||||||||||||||
Db   242  GCTGACGGAGCCCGGGCTGAGCGTGATCGTTGAGCCCCGTTGCCTCACCGACGCCAGGA 301

Qy   594  GAAGGACCACCATCAACGAGGAGCCGCCCGAGTGGGACCCGCGTCAGGCCGG 653
               |||||||||||||||||||||||||||||||||||||||||||||
Db   302  CCAGCGCCATGAAGAAACTACACCATCTACCGCGTGTCCATCACACATCCGCGT 361

Qy   654  CCCCGTGGCCCCCCGGCCCAGATGCGCGACCCCCGGCAGGCGACGACATCGCCGCACCAG 713
               |||||||||||||||||||||||||||||||||||||||||||
Db   362  CTTCCTCGCTCGTCGCGGCTGGTCTGGAAGTTCGACTTCGCGCCCTTCATGGTGCTCATCAG 421

Qy   714  CACCCTGCGAGGAGCAGGAGATCGCCTGGATGACCAGCAACCCCCGTCCCGTGGGCGGCAT 773
               |||||||||||||||||||||||||||||||||||||||||||||||
Db   362  CTTCCTCGCTCGTCGCGGCTGGTCTGGAAGTTCGACTTCGCGCCCTTCATGGTGCTCATCAG 421

```
COMMENT      Contact: Cathy Whitelaw
             Other GSSs: OG2BH19TH
             Consortium for Maize Genomics
             Unpublished (2002)
JOURNAL      
TITLE        
AUTHORS      Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
             Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
             Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.,
             1 (bases 1 to 883)
REFERENCE    
             clade; Panicoideae; Andropogoneae; Zea.
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
ORGANISM     Zea mays
SOURCE       Zea mays
KEYWORDS     GSS.
VERSION      CG274385.1  GI:34186526
ACCESSION    CG274385
             genomic survey sequence.
DEFINITION   OG2BH19TV ZM_0.7_1.5_KB zea mays genomic clone ZMMBMa075zC14,
LOCUS        CG274385                883 bp    DNA    linear    GSS 25-AUG-2003
RESULT 15
```

```
FEATURES             Location/Qualifiers
     source          1..883
                     /organism="zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa075zC14"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN
```

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

```
Query Match   5.0%;  Score 75.4;  DB 29;  Length 883;
Best Local Similarity 47.0%;  Pred. No. 8.8;
Matches 303;  Conservative 0;  Mismatches 336;  Indels 6;  Gaps 2;
```

Sequence alignment (Qy = Query, Db = Database):

```
Qy  659 TGGCCCCGGCCGAGATGCGCGGACCCCCGCCGCAGCGGCAGCATGCCCGCGGCCACCGACCC 718
        |||||||
Db   15 TGGCCCAGGCGCCGACGTGCGAGGACGCCGGCGTGCGCGGGGTGCTCGGGCCGGGGCGGTACC 74

Qy  719 TGCAGGAGCAGATCCGGCCTGGATGACCAGCAACCCCCCCCGGTGCCCGGGGCGGACATCTACA 778
        ||||||||
Db   75 CGCTGCGCGTTGCTCCCCGGGGATGCTGGAGGAAGTTCCCCAACCCCGGTGGACGCGCGGGT 134

Qy  779 AGCGGTGGATCATCCTGGGGCTTGAACAAGATCGTCGCGGATGTACAGCCCCGTGAGCATCC 838
        ||||||
Db  135 GGTGGCCGGACCCGCGGCGCGCAGCGCAGGCCAAGAGGCGTGGACGGCGTTCCCCGGTCCAGGA 194

Qy  779 ... 
```

(Alignment continues through multiple query/database segment blocks)